

<<数据挖掘与R语言>>

图书基本信息

书名：<<数据挖掘与R语言>>

13位ISBN编号：9787111407003

10位ISBN编号：7111407008

出版时间：2013-4

出版时间：机械工业出版社

作者：(葡)Luis Torgo

译者：李洪成,陈道轮,吴立明

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>

<<数据挖掘与R语言>>

内容概要

“如果你想学习如何用一款统计专家和数据挖掘专家所开发的免费软件包，那就选这本书吧。本书包括大量实际案例，它们充分体现了R软件的广度和深度。

”

——Bernhard Pfahringer, 新西兰怀卡托大学

本书利用大量给出必要步骤、代码和数据的具体案例，详细描述了数据挖掘的主要过程和技术，广泛涵盖数据大小、数据类型、分析目标、分析工具等方面的各种具有挑战性的问题。

本书的支持网站（<http://www.liaad.up.pt/~ltorgo/DataMiningWithR/>）给出了案例研究的所有代码、数据集以及R函数包。

本书特色

通过仔细选择的案例涵盖了主要的数据挖掘技术。

给出的代码和方法可以方便地复制或者改编后应用于自己的问题。

不要求读者具有R、数据挖掘或统计技术的基础知识。

包含R和MySQL基础知识的简介。

提供了对数据挖掘技术的特性、缺点和分析目标的基本理解。

<<数据挖掘与R语言>>

作者简介

Lu í s Torgo

葡萄牙波尔图大学计算机科学系副教授，现在在LIAAD实验室从事研究工作。
他是APPIA会员，同时还是OBEGEF的创办会员。

<<数据挖掘与R语言>>

书籍目录

出版者的话

推荐序

中文版序

译者序

前言

致谢

第1章简介

1.1如何阅读本书

1.2R简介

1.2.1R起步

1.2.2R对象

1.2.3向量

1.2.4向量化

1.2.5因子

1.2.6生成序列

1.2.7数据子集

1.2.8矩阵和数组

1.2.9列表

1.2.10数据框

1.2.11构建新函数

1.2.12对象、类和方法

1.2.13管理R会话

1.3MySQL简介

第2章预测海藻数量

2.1问题描述与目标

2.2数据说明

2.3数据加载到R

2.4数据可视化和摘要

2.5数据缺失

2.5.1将缺失部分剔除

2.5.2用最高频率值来填补缺失值

2.5.3通过变量的相关关系来填补缺失值

2.5.4通过探索案例之间的相似性来填补缺失值

2.6获取预测模型

2.6.1多元线性回归

2.6.2回归树

2.7模型的评价和选择

2.8预测7类海藻的频率

2.9小结

第3章预测股票市场收益

3.1问题描述与目标

3.2可用的数据

3.2.1在R中处理与时间有关的数据

3.2.2从CSV文件读取数据

3.2.3从网站上获取数据

<<数据挖掘与R语言>>

- 3.2.4从MySQL数据库读取数据
- 3.3定义预测任务
 - 3.3.1预测什么
 - 3.3.2预测变量是什么
 - 3.3.3预测任务
 - 3.3.4模型评价准则
- 3.4预测模型
 - 3.4.1如何应用训练集数据来建模
 - 3.4.2建模工具
- 3.5从预测到实践
 - 3.5.1如何应用预测模型
 - 3.5.2与交易相关的评价准则
 - 3.5.3模型集成：仿真交易
- 3.6模型评价和选择
 - 3.6.1蒙特卡罗估计
 - 3.6.2实验比较
 - 3.6.3结果分析
- 3.7交易系统
 - 3.7.1评估最终测试数据
 - 3.7.2在线交易系统
- 3.8小结
- 第4章侦测欺诈交易
 - 4.1问题描述与目标
 - 4.2可用的数据
 - 4.2.1加载数据至R
 - 4.2.2探索数据集
 - 4.2.3数据问题
 - 4.3定义数据挖掘任务
 - 4.3.1问题的不同解决方法
 - 4.3.2评价准则
 - 4.3.3实验方法
 - 4.4计算离群值的排序
 - 4.4.1无监督方法
 - 4.4.2有监督方法
 - 4.4.3半监督方法
 - 4.5小结
- 第5章微阵列样本分类
 - 5.1问题描述与目标
 - 5.1.1微阵列实验背景简介
 - 5.1.2数据集ALL
 - 5.2可用的数据
 - 5.3基因（特征）选择
 - 5.3.1基于分布特征的简单过滤方法
 - 5.3.2ANOVA过滤
 - 5.3.3用随机森林进行过滤
 - 5.3.4用特征聚类的组合进行过滤
 - 5.4遗传学异常的预测

<<数据挖掘与R语言>>

- 5.4.1 定义预测任务
- 5.4.2 模型评价标准
- 5.4.3 实验过程
- 5.4.4 建模技术
- 5.4.5 模型比较
- 5.5 小结
- 参考文献
- 主题索引
- 数据挖掘术语索引
- R函数索引

<<数据挖掘与R语言>>

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>