

<<数据挖掘>>

图书基本信息

书名：<<数据挖掘>>

13位ISBN编号：9787111172482

10位ISBN编号：7111172485

出版时间：2005-9

出版时间：机械工业出版社

作者：威滕

页数：524

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>>

内容概要

本书对1999年的初版做了重大的改动。

虽说核心概念没有变化，但本书进行了更新使其能反映过去5年里的变化，参考文献几乎翻了一番。新版的重要部分包括：30个新的技术章节；一个加强了具有交互式界面的Weka机器学习工作平台；有关神经网络的完整信息，一个有关贝叶斯网络的新节；等等。

本书提供了机器学习概念的完整基础，此外还针对实际工作中应用相关工具和技术提出了一些建议，在本书中你将发现：
成功数据挖掘技术的核心算法，包括历经考验的真实技术及前沿的方法。

转换输入或输出以改善性能的方法。

可下载的Weka软件??一个用于数据挖掘任务的机器学习算法的集合，包括用于数据预处理、分类、回归、聚类、关联规则以及在新的交互式界面上可视化的工具。

作者简介

Lan H.Witten新西兰怀卡托大学计算机科学系教授，ACM和新西兰皇家学会成员。
他曾荣获2004年国际信息处理研究联合会颁发的Namur奖项，这是一个两年一度的荣誉奖项，用于奖励那些在信息和通信技术的社会应用方面做出杰现贡献及具有国际影响的人。
他的著作包括《Managing Gi

书籍目录

Foreword Preface Part I Machine learning tools and techniques 1. What's it all about? 1.1 Data mining and machine learning 1.2 Simple examples: the weather problem and others 1.3 Fielded applications 1.4 Machine learning and statistics 1.5 Generalization as search 1.6 Data mining and ethics 1.7 Further reading 2. Input: Concepts, instances, attributes 2.1 What's a concept? 2.2 What's in an example? 2.3 What's in an attribute? 2.4 Preparing the input 2.5 Further reading 3. Output: Knowledge representation 3.1 Decision tables 3.2 Decision trees 3.3 Classification rules 3.4 Association rules 3.5 Rules with exceptions 3.6 Rules involving relations 3.7 Trees for numeric prediction 3.8 Instance-based representation 3.9 Clusters 3.10 Further reading 4. Algorithms: The basic methods 4.1 Inferring rudimentary rules 4.2 Statistical modeling 4.3 Divide-and-conquer: constructing decision trees 4.4 Covering algorithms: constructing rules 4.5 Mining association rules 4.6 Linear models 4.7 Instance-based learning 4.8 Clustering 4.9 Further reading 5. Credibility: Evaluating what's been learned 5.1 Training and testing 5.2 Predicting performance 5.3 Cross-validation 5.4 Other estimates 5.5 Comparing data mining schemes 5.6 Predicting probabilities 5.7 Counting the cost 5.8 Evaluating numeric prediction 5.9 The minimum description length principle 5.10 Applying MDL to clustering 5.11 Further reading 6. Implementations: Real machine learning schemes 6.1 Decision trees 6.2 Classification rules 6.3 Extending linear models 6.4 Instance-based learning 6.5 Numeric prediction 6.6 Clustering 6.7 Bayesian networks 7. Transformations: Engineering the input and output 7.1 Attribute selection 7.2 Discretizing numeric attributes 7.3 Some useful transformations 7.4 Automatic data cleansing 7.5 Combining multiple models 7.6 Using unlabeled data 7.7 Further reading 8. Moving on: Extensions and applications 8.1 Learning from massive datasets 8.2 Incorporating domain knowledge 8.3 Text and Web mining 8.4 Adversarial situations 8.5 Ubiquitous data mining 8.6 Further reading Part II: The Weka machine learning workbench 9. Introduction to Weka 9.1 What's in Weka? 9.2 How do you use it? 9.3 What else can you do? 9.4 How do you get it? 10. The Explorer 10.1 Getting started 10.2 Exploring the Explorer 10.3 Filtering algorithms 10.4 Learning algorithms 10.5 Meta-learning algorithms 10.6 Clustering algorithms 10.7 Association-rule learners 10.8 Attribute selection 11. The Knowledge Flow interface 11.1 Getting started 11.2 Knowledge Flow components 11.3 Configuring and connecting the components 11.4 Incremental learning 12. The Experimenter 12.1 Getting started 12.2 Simple setup 12.3 Advanced setup 12.4 The Analyze panel 12.5 Distributing processing over several machines 13. The command-line interface 13.1 Getting started 13.2 The structure of Weka 13.3 Command-line options 14. Embedded machine learning 15. Writing new learning schemes References Index

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>