

<<数据挖掘>>

图书基本信息

书名：<<数据挖掘>>

13位ISBN编号：9787111127697

10位ISBN编号：7111127692

出版时间：2003-9

出版时间：机械工业出版社

作者：Ian H.Witten,Eibe Frank

页数：369

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

这是一本将数据挖掘算法和数据挖掘实践完美结合起来的优秀教材。作者以其丰富的经验，对数据挖掘的概念和数据挖掘所有的技术（特别是机器学习）进行了深入浅出的介绍，并对应用机器学习工具进行数据挖掘给出了良好的建议。数据挖掘中的各个关键要素也事例融合在众多实例中加以介绍。

本书还介绍了Weka这种基于Java的软件系统。该软件系统可以用来分析数据集，找到适用的模式，进行正确的分析，也可以用来开发自己的机器学习方案。

本书的主要特点：

解释数据挖掘算法的原理。

通过实例帮助读者根据实际情况选择合适的算法，并比较和评估不同方法得出的结果。

介绍提高性能的技术，包括数据处理以及组合不同方法得到的输出。

提供了本书所有的Weka软件和附加学习材料，可以从<http://www.mkp.com/datamining>上下载这些资料。

<<数据挖掘>>

作者简介

Lan H.Witten，新西兰怀卡托大学计算机科学系教授。

他是ACM和新西兰皇家学会的成员，并参加了英国、美国、加拿大和新西兰的专业计算、信息检索、工程等协会。

他著有多部著作，是多家技术杂志的作者，发表过大量论文。

## 书籍目录

Foreword Preface 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 application 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 Further reading 5 Credibility: Evaluation 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 Implementation: Real machine learning schemes 6.1 Decision trees 6.2 Classification rules 6.3 Extending linear classification: Support vector machines 6.4 Instance-based learning 6.5 Numeric prediction 6.6 Clustering 7 Moving on: Engineering the input and output 7.1 Attribute selection 7.2 Discretizing numeric attributes 7.3 Automatic data cleansing 7.4 Combining multiple models 7.5 Further reading 8 Nuts and bolts: Machine learning algorithms in Java 8.1 Getting started 8.2 Javadoc and the class library 8.3 Processing dataset using the machine learning programs 8.4 Embedded machine learning 8.5 Writing new learning schemes 9 Looking forward 9.1 learning from massive datasets 9.2 Visualizing machine learning 9.3 Incorporation domain knowledge 9.4 Text mining 9.5 Mining the World Wide Web 9.6 Further reading References Index About the authors

<<数据挖掘>>

媒体关注与评论

书评本书是综合运用数据挖掘、数据分析、信息理论通讯机器学习技术的里程碑。

<<数据挖掘>>

编辑推荐

其它版本请见：《经典原版书库&middot;数据挖掘：实用机器学习技术（英文版）（第2版）（新版）》

#### 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>