<<数据挖掘导论>>

图书基本信息

书名：<<数据挖掘导论>>

13位ISBN编号：9787111316701

10位ISBN编号：7111316703

出版时间：2010.9

出版时间：机械工业出版社

作者：(美)Pang-Ning Tan,Michael Steinbach,Vipin Kumar

页数：769

前言

Advances in data generation and collection are producing data sets of massire size in commerce and a variety of scientific disciplines. Data warehouses store details of the sales and operations of businesses, Earth-orbiting satelfites beam high-resolution images and sensor data back to Earth. and genomics ex- periments generate sequence, structural, and functional data for an increasing number of organisms. The ease with Which data can now be gathered and stored has created a new attitude toward data analysis：Gather whatever data you can whenever and wherever possible. It has become an article of faith that the gathered data will have value. either for the purpose that initially motivated its collection or for purposes not yet envisioned. The field of data mining grew out of the limitations of current data analysis techniques in handling the challenges posed by these new types of data sets. Data mining does not replace other areas of data analysis, but rat. Her takes them as the foundation for much of its work. While some areas of data mining, such as association analysis, are unique to the field, other areas, such as clustering, classification, and anomaly detection, build upon a long history of work on these topics in other fields. Indeed. the willingness of data mining researchers to draw upon existing techniques has contributed to the strength and breadth of the field, as well as to its rapid growth.

<<数据挖掘导论>>

内容概要

本书全面介绍了数据挖掘的理论和方法，着重介绍如何用数据挖掘知识解决各种实际问题，涉及学科领域众多，适用面广。
书中涵盖5个主题：数据、分类、关联分析、聚类和异常检测。
除异常检测外，每个主题都包含两章：前面一章讲述基本概念、代表性算法和评估技术，后面一章较深入地讨论高级概念和算法。
目的是使读者在透彻地理解数据挖掘基础的同时，还能了解更多重要的高级主题。
.包含大量的图表、综合示例和丰富的习题。
·不需要数据库背景。
只需要很少的统计学或数学背景知识。
·网上配套教辅资源丰富，包括PPT、习题解答、数据集等。

<<数据挖掘导论>>

作者简介

作者：（美国）谭（Pang-Ning Tan） （美国）斯坦巴克（Michael Steinbach） （美国）库马尔（Vipin Kumar）Pang.Ning Tan现为密歇根州立大学计算机与工程系助理教授，主要教授数据挖掘、数据库系统等课程。
他的研究主要关注于为广泛的应用（包括医学信息学、地球科学、社会网络、Web挖掘和计算机安全）开发适用的数据挖掘算法。
Michael Steinbach拥有明尼苏达大学数学学士学位、统计学硕士学位和计算机科学博士学位，现为明尼苏达大学双城分校计算机科学与工程系助理研究员。
Vipin Kumar现为明尼苏达大学计算机科学与工程系主任和William Norris教授。
1 988年至2005年。
他曾担任美国陆军高性能计算研究中心主任。

<<数据挖掘导论>>

书籍目录

章节摘录

插图：What Is an attribute?We start with a more detailed definition of an attribute.Definition 2.1. An attribute is a property or characteristic of an object that may vary, either from one object to another or from one time to another.For example, eye color varies from person to person, while the temperature of an object varies over time. Note that eye color is a symbolic attribute with a small number of possible values brown, black, blue, green, hazel, etc.}, while temperature is a numerical attribute with a potentially unlimited number of values.At the most basic level, attributes are not about numbers or symbols. However, to discuss and more precisely analyze the characteristics of objects, we assign numbers or symbols to them. To do this in a well-defined way, we need a measurement scale. Definition 2.2. A measurement scale is a rule (function) that associates a numerical or symbolic value with an attribute of an object.Formally, the process of measurement is the application of a measurement scale to associate a value with a particular attribute of a specific object. While this may seem a bit abstract, we engage in the process of measurement all the time. For instance, we step on a bathroom scale to determine our weight, we classify someone as male or female, or we count the number of chairs in a room to see if there will be enough to seat all the people coming to a meeting. In all these cases, the "physical value" of an attribute of an object is mapped to a numerical or symbolic value.With this background, we can now discuss the type of an attribute, a concept that is important in determining if a particular data analysis technique is consistent with a specific type of attribute.

编辑推荐

《数据挖掘导论(英文版)》是经典原版书库。

版权说明