《数据算法》

图书基本信息

书名：《数据算法》

13位ISBN编号：9787512395949

出版时间：2016-10-1

作者：Mahmoud Parsian

译者：苏金国,杨健康

《数据算法》

内容概要

《数据算法：Hadoop/Spark大数据处理技巧》介绍了很多基本设计模式、优化技术和数据挖掘及机器学习解决方案，以解决生物信息学、基因组学、统计和社交网络分析等领域的很多问题。这还概要介绍了MapReduce、Hadoop和Spark。
主要内容包括：
■ 完成超大量交易的购物篮分析。
■ 数据挖掘算法（K-均值、KNN和朴素贝叶斯）。
■ 使用超大基因组数据完成DNA和RNA测序。
■ 朴素贝叶斯定理和马尔可夫链实现数据和市场预测。
■ 推荐算法和成对文档相似性。
■ 线性回归、Cox回归和皮尔逊（Pearson）相关分析。
■ 等位基因频率和DNA挖掘。
■ 社交网络分析（推荐系统、三角形计数和情感分析）。

《数据算法》

作者简介

Mahmoud Parsian，计算机科学博士，是一位热衷于实践的软件专家，作为开发人员、设计人员、架构师和作者，他有30多年的软件开发经验。目前领导着Illumina的大数据团队，在过去15年间，他主要从事Java (服务器端)、数据库、MapReduce和分布式计算的有关工作。Mahmoud还著有《JDBC Recipes》和《JDBC Metadata， MySQL，and Oracle Recipes》等书（均由Apress出版）。

《数据算法》

书籍目录

《数据算法》

版权说明