数据科学实战

数据科学实战_下载链接1_

著者:[美] Rachel Schutt

出版者:人民邮电出版社

出版时间:2015-3

装帧:平装

isbn:9787115383495

- 统计推断、探索性数据分析（EDA）及数据科学工作流程
- 算法

- 垃圾邮件过滤、朴素贝叶斯和数据清理
- 逻辑回归
- 金融建模
- 推荐引擎和因果关系
- 数据可视化
- 社交网络与数据新闻
- 数据工程、MapReduce、Pregel和Hadoop

作者介绍:

作者简介：

Rachel Schutt

美国新闻集团旗下数据科学部门高级副总裁、哥伦比亚大学统计系兼职教授、约翰逊实验室高级研究科学家，同时也是哥伦比亚大学数据科学及工程研究所教育委员会的发起人之一。她曾在谷歌研究院工作数年，负责设计算法原型并通过建模理解用户行为。

Cathy O'Neil

约翰逊实验室高级数据科学家、哈佛大学数学博士、麻省理工学院数学系博士后、巴纳德学院教授，曾发表过大量算术代数几何方面的论文。他曾在著名的全球投资管理公司D.E.
Shaw担任对冲基金金融师，后加入专门评估银行和对冲基金风险的软件公司RiskMetrics，个人博客：mathbabe.org。

译者简介：

冯凌秉

澳大利亚国立大学统计学博士，本科和研究生分别毕业于中南财经政法大学和中国人民大学。现在，他任职于江西财经大学金融管理国际研究院，任讲师、硕士生导师，研究方向为应用统计与金融计量。

王群锋

毕业于西安电子科技大学，现任职于IBM西安研发中心，从事下一代统计预测软件的开发运维工作。

数据科学实战_下载链接1_

标签

数据科学

数据分析

数据挖掘

机器学习

大数据

统计

计算机

数据

评论

很不错的一本湿货，翻译好的没话说，连“无厘头”都被翻出来了，很想知道原文是不是nonsense。。。以后还会翻看里面的R程序

翻译还算通顺，但是错别字太多了。里面有一些东西还是能给人启示的。

對於ML 來說，這本書講得太淺。興許對於DS 還尚可。不過最後對於data scientist 所應具備的技能，還是蠻認可的。scientist 比engineer
做好了還是技高一籌，對於研究結果，論文輸出或報告輸出擴大影響還是蠻有好處。只是本書翻譯欠佳，讀來不禁搖頭嘆息。要是要理解各算法，還是MMDS 更值得推薦。

译者也太能放飞自我了。

意外地不错，更多是思维方式。干货不少，比如欺诈侦测那一节，虽然只是大概描述了一下square的风险管理模型的架构，但是因为我自己做这一块，所以就知道所言非虚。

涵盖面很广，个别章节还是有一定难度，给出的代码可操作性不高。整本书更像是一系列课程教案的汇总，含金量还是不错的，学到了一些知识。

这本书涉及的面包含数据科学的各个层面，是相关专家客座课堂讲义的分析和总结。相对比较烧脑的是算法的原理解释与过程，从通用的分类、聚类算法的，k近邻、k均值，朴素贝叶斯，决策树，到应用在金融、推荐系统等业务领域的案例与分析，深浅适宜，不过依然对数学一般的我产生不小压力。

特别好，给了结实的框架，接下来想要探究什么就可以按图索骥了。书里给出的详细的参考书目几乎每一本都是在别的地方交叉推荐过的，含金量很高。（处在入门阶段，非数学专业）

结合一线数据科学家日常工作的一手资料，在山顶眺望全局，非工具书，也有不明觉厉之处，但对打开视野大有裨益，技能和方法岂是朝夕之间，实战嘛，纸上得来终觉浅

强调数据科学总体视角，区别于算法教材，不难但适合有一定基础后阅读，有很多大牛的经验

优秀!
基于哥大的《数据科学导论》讲义，作者想挑战的是"数据科学是不可能在大学里或者书本里学会的"这一论断。作者并未对"数据科学"下精确定义，而是通过"数据科学家在做的事"来定义"数据科学"。"师者，所以传道受业解惑也"。授人以鱼不如授人以渔。
身为新兴和交叉学科，全书通过从业者本身在做的事，给出了"数据科学"的各个面向。不同学科背景的人，可以结合自身优势和兴趣点，按图索骥，找到适合自身的发展路径。"我希望他们能从中得到激励，或者学到一些有用的工具，来让这个世界变得更好，而不是更坏"。
发展初期，也即规则制定期，学的开心，玩的尽兴。"不必现在就规划好未来，走点弯路也没什么，谁知道这一路上你会发现什么呢?"

师者，所以传道授业解惑也。这本书做到了

看后依旧手无寸铁，却隐约有了去挖矿打铁器的动力。随便摘抄一段：下一代数据科学家会怎么做？1.
对一切保持怀疑态度：怀疑模型本身，模型在什么情况下会失败，如何使用，因何会被无用；2. 认识到「模型的反馈循环与潜在模型之间的博弈」

我的数据科学开蒙。

数据驱动的政治分析

什么是data science

内容比较简单。但是一些平时不会注意到的细节的讨论还是不错的。

对具体算法细节未做深入探讨，但属于名副其实的实战，值得一读，尤其推荐六九两章的部分内容。

主要指明大数据的应用领域, 偏应用范畴, 对Data Scientist来说很赞...

一翻就不明觉厉看不下去的类型，拜服

数据科学实战_下载链接1_

书评

这本书蛮不错的，就是看的时候碰到一些小错误，记录如下，如果本书的编者看到了，也方便勘误。 P43 第11行 “事” 改为 “是” P45 第9行 “歌” 改为 “个” P52 图3-6说明文字第2行 “直” 改为 “致” P96 正文第6行 “Emprical” 改为 “Empirical” P103 倒数第4行 “...

很喜欢此书，但首先要说这本书不是用来入门算法看的。 data science的方法是各种统计学计算机方法的综合，所以所有对统计学有较好的数理基础，对各种统计推断方法或数据挖掘算法有较好理解的童鞋可以通过翻阅此书，从各个角度打开对data science的认知。如果没有很好的相关知...

Now that answering complex and compelling questions with data can make the difference in an election or a business model, data science is an attractive discipline. But how can you learn this wide-ranging, interdisciplinary field? With this book, you’ll get...

我看过了 我看过了

我看过了 我看...

数据科学实战_下载链接1_