

人人可懂的数据科学



[人人可懂的数据科学_下载链接1](#)

著者:[英] 约翰·D·巴罗

出版者:

出版时间:

装帧:

isbn:9787111637264

数据科学的主要目标就是通过数据分析来改进决策，它与数据挖掘、机器学习等领域紧密相关，但范围更广。本书简要介绍了该领域的发展、基础知识，并阐释了数据科学项

目的各个阶段。书中既考虑数据基础架构和集成多个数据源数据所面临的挑战，又介绍机器学习基础并探讨如何应用机器学习专业技术解决现实问题。还综述了伦理和法律问题、数据法规的发展以及保护隐私的计算方法。最后探讨了数据科学的未来影响，并给出数据科学项目成功的原则。

作者介绍:

约翰·D.凯莱赫(John D. Kelleher) 布伦丹·蒂尔尼(Brendan Tierney) 著：约翰·D.凯莱赫(John D. Kelleher) 是都柏林理工学院计算机科学学院的教授以及信息、通信和娱乐研究所的学术负责人。他的研究得到了ADAPT中心的支持，该中心由爱尔兰科学基金会 (Grant 13 / RC / 2106) 资助，同时也接受欧洲区域发展基金的资助。他还是《Fundamentals of Machine Learning for Predictive Data Analytics》的作者之一。

布伦丹·蒂尔尼(Brendan Tierney) 是都柏林理工学院计算机科学学院的讲师，同时也是Oracle ACE 主任，还著有多本基于Oracle技术的数据挖掘类著作。

目录: 译者序

前言

致谢

作者简介

第1章 什么是数据科学 1

1.1 数据科学简史 5

1.1.1 数据收集简史 5

1.1.2 数据分析简史 9

1.1.3 数据科学的产生与发展 14

1.2 数据科学用于何处 20

1.2.1 销售和营销中的数据科学 21

1.2.2 数据科学在政府中的应用 22

1.2.3 数据科学在竞技体育中的应用 23

1.3 为什么是现在 25

1.4 关于数据科学的神话 28

第2章 什么是数据，什么是数据集 31

2.1 关于数据的观点 38

2.2 数据可以积累，而智慧不能 43

2.3 CRISP-DM 45

第3章 数据科学生态系统 54

3.1 将算法迁移至数据 61

3.1.1 传统数据库与现代的传统数据库 64

3.1.2 大数据架构 67

3.1.3 混合数据库世界 69

3.2 数据准备和集成 72

第4章 机器学习 77

4.1 有监督学习与无监督学习 78

4.2 学习预测模型 83

4.2.1 相关性不等同于因果，但它有时非常有用 84

4.2.2 线性回归 90

4.2.3 神经网络与深度学习 96

4.2.4 决策树 108

4.3 数据科学中的偏差 114

4.4 评估模型：泛化而不是记忆 116

4.5 摘要	119
第5章 标准的数据科学任务	121
5.1 谁是我们的目标客户（聚类）	122
5.2 这是欺诈吗（异常值检测）	128
5.3 你要配份炸薯条吗（关联规则挖掘）	131
5.4 流失还是不流失，这是一个问题（分类）	136
5.5 它价值几何（回归）	141
第6章 隐私与道德	143
6.1 商业利益与个人隐私	145
6.1.1 数据科学的道德启示：画像与歧视	148
6.1.2 数据科学的道德含义：创建一个全景监狱	154
6.2 隐私保护	157
6.2.1 保护隐私的计算方法	159
6.2.2 规范数据使用和保护隐私的法律框架	161
6.3 通往道德的数据科学之路	164
第7章 未来趋势与成功准则	172
7.1 医疗数据科学	172
7.2 智慧城市	174
7.3 数据科学项目准则：为什么会成功或失败	177
7.4 终极思考	185
术语表	188
延伸阅读	201
参考文献	203
• • • • •	(收起)

[人人易懂的数据科学 下载链接1](#)

标签

数据科学

评论

边读边向身边人推荐，全面但又篇幅很小，详略得当。第七章的数据科学成功准则，简直字字印证了当前工作的现状。好书让人激动啊！

Low A-

作为小白看完很有收获，译者大大辛苦啦

[人人易懂的数据科学_下载链接1](#)

书评

[人人易懂的数据科学_下载链接1](#)