



北京邮电大学

BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS



Data Analytics with R
语言数据分析



格言联璧话学习

艾新波 / 2018 • 北京



课程体系



R语言数据分析



上部：论道



- 第1章 气象万千、数以等观
- 第2章 所谓学习、归类而已
- 第3章 格言联璧话学习
- 第4章 源于数学、归于工程



中部：执具



- 第5章 工欲善其事必先利其器
- 第6章 基础编程
- 第7章 数据对象

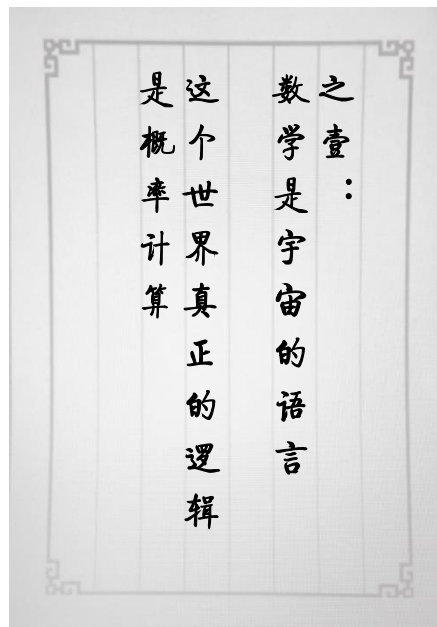
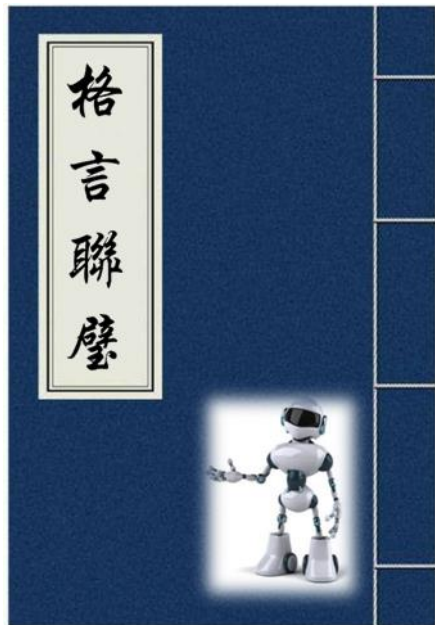


下部 博术



- 第8章 人人都爱tidyverse
- 第9章 最美不过数据框
- 第10章 观数以形
- 第11章 相随相伴、谓之关联
- 第12章 既是世间法、自当有分别
- 第13章 方以类聚、物以群分
- 第14章 庐山烟雨浙江潮

机器学习：格言联璧



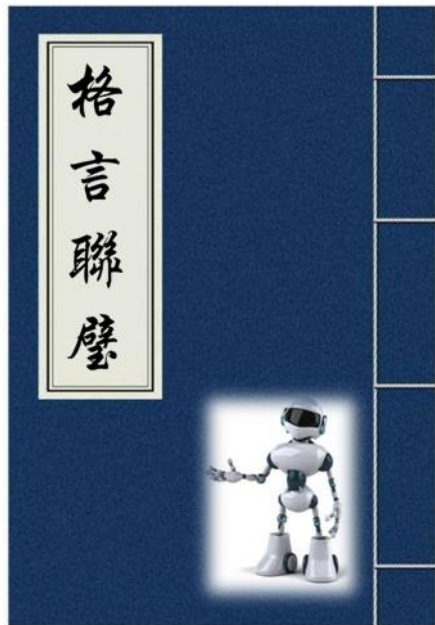
机器学习：格言联璧

格言联璧



之貳：
相关并不意味着因果
相关关系、函数关系
都不能等同于因果关
系，但这无损于他们
的价值

机器学习：格言联璧



之参：
所有的模型都是错的
但是有些是有用的
模型的评估与模型的
建立同等重要

机器学习：格言联璧

格言联璧

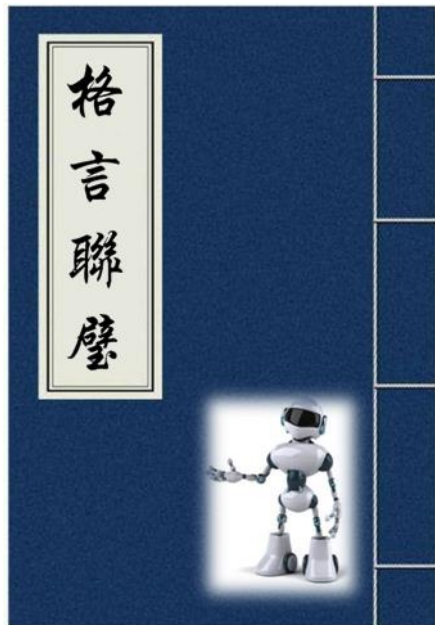


x_1	x_2	y
19.99	36.07	123.95
54.97	5.93	119.11
7.19	75.9	96.89
2.06	66.03	111.9
20.82	7.19	151.98
41.05	43.08	95.87
21.04	40.93	118.02
43.11	7.85	129.05
34.9	2.91	142.21
42.03	0.91	137.08
24.89	56.11	98.99
39.17	39.04	101.77

No Free Lunch

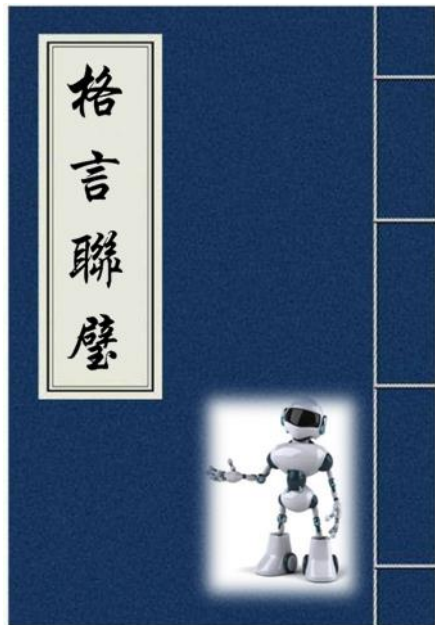
之肆：
没有明确证据表明一个
模型会优于另一个模型
模型的选择取决于任务
情境

机器学习：格言联璧



之伍：
成功的机器学习应用
不是拥有最好的算法
而是拥有最多的数据

机器学习：格言联璧



机器学习：格言联璧

格言联璧



之染：
数据和特征决定了机
器学习的上限
而模型和算法只是逼
近这个上限而已

机器学习：格言联璧

格言联璧



之捌：
如何赢得机器学习的
竞赛？
就是把别人的模型拿
过来然后集成在一起

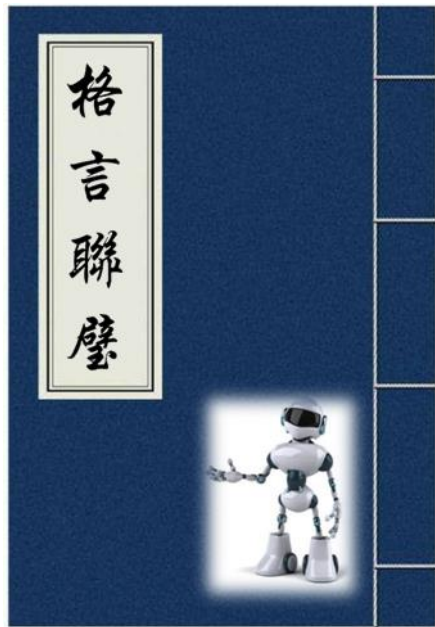
机器学习：格言联璧

格言联璧



之致：
数据科学家是一个比
计算机科学家懂更多计
统计学
比统计学家懂更多计
算机科学的人

机器学习：格言联璧



A decorative blue border with rounded corners frames the entire slide. Two thin blue crosshair-like lines are positioned diagonally, one in the upper right and one in the lower left, intersecting at the center of the text.

谢谢聆听
Thank you

教师个人联系方式

艾新波

手机: 13641159546

QQ: 23127789

微信: 13641159546

E-mail: 13641159546@126.com

axb@bupt.edu.cn

地址: 北京邮电大学科研楼917室

课程网址: <https://github.com/byaxb/RDataAnalytics>

