

R语言预测实战



[R语言预测实战 下载链接1](#)

著者:著

出版者:电子工业出版社

出版时间:2016-10

装帧:平装

isbn:9787121298547

R语言具有上手快、效率高的特点，它横跨金融、生物、医学、互联网等多个领域，主要用于统计、建模及可视化。预测是数据挖掘的主要作用之一，也是大数据时代的核心价值所在。基于R语言来做预测，兼具效率和价值的双重属性，这是《R语言预测实战

》的一大亮点。

《R语言预测实战》共分为三部分。第一部分讲预测基础，主要涵盖预测概念理解、预测方法论、分析方法、特征技术、模型优化及评价，读者通过这部分内容的学习，可以掌握进行预测的基本步骤和方法思路；第二部分讲预测算法，该部分包含了多元回归分析、复杂回归分析、时间序列及进阶算法，内容比较有难度，需要细心体会；第三部分讲预测案例，包括短期日负荷曲线预测和股票价格预测两个实例，读者可以了解到实施预测时需要关注的技术细节。

希望读者在看完《R语言预测实战》后，能够将本书的精要融会贯通，进一步在工作和学习实践中提炼价值。

作者介绍:

游皓麟，高级数据分析师，目前专注于NLP、知识图谱以及深度学习的研究与实现。曾服务于华为技术软件有限公司等企业，多次出席R语言会议并发表演讲，在小象学院担任过R语言数据挖掘和机器学习讲师。

目录: 第一部分 预测入门篇

第1章 预测入门 2

1.1 什么是预测 3

1.1.1 预测的定义 3

1.1.2 预测的特点 4

1.1.3 预测的分类 4

1.1.4 预测的基本原则 5

1.2 大数据与预测 7

1.2.1 什么是大数据 7

1.2.2 大数据预测的优势 9

1.2.3 大数据预测的特征 9

1.2.4 大数据预测案例 12

1.3 预测利器之R语言 22

1.3.1 R语言简介 22

1.3.2 R语言预测初步 33

1.3.3 R语言预测常用包 42

第2章 预测方法论 44

2.1 预测流程 45

2.1.1 确定主题 45

2.1.2 收集数据 47

2.1.3 选择方法 49

2.1.4 分析规律 53

2.1.5 建立模型 57

2.1.6 评估效果 60

2.1.7 发布模型 60

2.2 指导原则 61

2.2.1 界定问题 61

2.2.2 判断预测法 63

2.2.3 外推预测法 64

2.2.4 因果预测法 66

2.3 团队构成 66

2.3.1 成员分类 67

2.3.2 数据氛围 68

2.3.3 团队合作 70

第3章 分析方法	72
3.1 相关分析	73
3.1.1 自相关分析	73
3.1.2 偏相关分析	74
3.1.3 简单相关分析	75
3.1.4 互相关分析	83
3.1.5 典型相关分析	84
3.2 对应分析	88
3.3 频谱分析	92
3.4 趋势分析	94
3.5 聚类分析	96
3.5.1 K-Means算法	96
3.5.2 系统聚类算法	99
3.6 关联分析	105
3.6.1 关联规则挖掘：Apriori与Eclat算法	105
3.6.2 序列模式挖掘：SPADE算法	119
第4章 特征构建技术	132
4.1 特征变换	133
4.1.1 概念分层	134
4.1.2 标准化	136
4.1.3 离散化	139
4.1.4 函数变换	142
4.1.5 深入表达	143
4.2 特征组合	143
4.2.1 基于特定的领域知识	143
4.2.2 二元组合	144
4.2.3 高阶多项式	148
4.3 自动生成：基于遗传编程的方法	152
4.3.1 基本思路	153
4.3.2 特征表达式	154
4.3.3 产生初始种群	162
4.3.4 计算适应度	163
4.3.5 选择、交叉和变异	165
4.3.6 实例分析	171
第5章 特征选择方法	176
5.1 直接法	177
5.2 单元法	177
5.2.1 Pearson相关系数	177
5.2.2 距离相关系数	179
5.2.3 单因素方差分析	181
5.2.4 信息增益	184
5.2.5 卡方检验	191
5.2.6 Gini系数	194
5.3 多元法	198
5.3.1 逐步回归	198
5.3.2 随机森林	203
5.3.3 遗传算法	210
第6章 模型参数优化	213
6.1 交叉验证	214
6.2 网格搜索	215
6.3 遗传算法	217
6.3.1 基本概念	217
6.3.2 遗传算法算例	218
6.3.3 遗传算法实现步骤	223

6.3.4 遗传算法R语言实现	223
6.3.5 R语言mcga和genalg包的应用	228
6.4 粒子群优化	233
6.4.1 基本概念及原理	233
6.4.2 粒子群算法R语言实现	235
6.4.3 粒子群算法实现步骤	238
6.4.4 R语言pso包的应用	239
6.5 模拟退火	241
6.5.1 基本概念及原理	241
6.5.2 模拟退火算法R语言实现	242
6.5.3 模拟退火算法实现步骤	244
6.5.4 R语言GenSA和stats包的应用	245
第7章 预测效果评估	250
7.1 概率预测评估方法	251
7.1.1 混淆矩阵	251
7.1.2 ROC曲线	255
7.1.3 KS曲线	261
7.1.4 累计收益图	263
7.1.5 累计提升图	264
7.1.6 累计响应图	266
7.2 数值预测评估方法	267
7.2.1 常见评估指标	267
7.2.2 ASD累计收益图	270
第二部分 预测算法篇	
第8章 线性回归及其优化	274
8.1 多元线性回归	275
8.1.1 回归模型和基本假定	275
8.1.2 最小二乘估计	276
8.1.3 回归方程和回归系数的显著性检验	276
8.1.4 多重共线性	277
8.2 Ridge回归	280
8.2.1 基本概念	281
8.2.2 岭迹曲线	281
8.2.3 基于GCV准则确定岭参数	283
8.2.4 Ridge回归的R语言实现	284
8.3 Lasso回归	285
8.3.1 基本概念	285
8.3.2 使用LAR求解Lasso	286
8.3.3 Lasso算法的R语言实现	288
8.3.4 R语言lars包的应用	290
8.4 分位数回归	292
8.4.1 基本概念	292
8.4.2 分位数回归的计算	294
8.4.3 用单纯形法求解分位数回归及R语言实现	296
8.4.4 R语言quantreg包的应用	298
8.5 稳健回归	300
8.5.1 基本概念	301
8.5.2 M-估计法及其R语言实现	301
8.5.3 应用R语言MASS包实现稳健回归	304
第9章 复杂回归分析	307
9.1 梯度提升回归树 (GBRT)	308
9.1.1 Boosting方法简介	308
9.1.2 AdaBoost算法	308
9.1.3 提升回归树算法	311

9.1.4 梯度提升	312
9.1.5 GBRT的R语言实现	314
9.1.6 R语言gbm包的应用	316
9.2 神经网络	320
9.2.1 基本概念	320
9.2.2 单层感知器学习算法	322
9.2.3 SLP回归算法的R语言实现	323
9.2.4 BP神经网络学习算法	325
9.2.5 BP回归算法的R语言实现	327
9.2.6 RBF神经网络学习算法	330
9.2.7 RBF回归算法的R语言实现	332
9.2.8 Elman神经网络学习算法	334
9.2.9 Elman回归算法的R语言实现	336
9.2.10 使用R语言包构建神经网络	338
9.3 支持向量机回归	343
9.3.1 基本问题	344
9.3.2 LS-SVMR算法	347
9.3.3 LS-SVMR算法的R语言实现	348
9.4 高斯过程回归	349
9.4.1 GPR算法	350
9.4.2 GPR算法的R语言实现	352
9.4.3 R语言kernlab包的应用	355
第10章 时间序列分析	358
10.1 Box-Jenkins方法	359
10.1.1 p阶自回归模型	359
10.1.2 q阶移动平均模型	361
10.1.3 自回归移动平均模型	363
10.1.4 ARIMA模型	365
10.1.5 ARIMA模型的R语言实现	367
10.1.6 R语言forecast包的应用	373
10.2 门限自回归模型	376
10.2.1 TAR模型的基本原理	376
10.2.2 TAR模型的R语言实现	377
10.2.3 R语言TSA包的应用	380
10.3 GARCH模型族	382
10.3.1 线性ARCH模型	382
10.3.2 GRACH模型	383
10.3.3 EGARCH模型	384
10.3.4 Power ARCH模型	384
10.3.5 PARCH模型的R语言实现	385
10.3.6 R语言fGarch包的应用	395
10.4 向量自回归模型	398
10.4.1 VAR模型基本原理	398
10.4.2 VAR模型的R语言实现	399
10.4.3 R语言vars包的应用	403
10.5 卡尔曼滤波器算法	405
10.5.1 Kalman滤波算法初步	406
10.5.2 Kalman滤波的R语言实现	407
10.5.3 R语言FKF包的应用	409
第三部分 预测应用篇	
第11章 短期日负荷曲线预测	414
11.1 电力行业负荷预测介绍	415
11.2 短期日负荷曲线预测的基本要求	415
11.3 预测建模准备	416

- 11.3.1 基础数据采集 416
- 11.3.2 缺失数据处理及平滑 418
- 11.3.3 潜在规律分析 421
- 11.4 基于RBF神经网络的预测 426
 - 11.4.1 RBF网络结构设计 426
 - 11.4.2 确定最优参数 427
 - 11.4.3 建模并实现预测 431
 - 11.4.4 效果评估 433
- 11.5 基于LS-SVMR算法的预测 435
 - 11.5.1 确定最优参数 436
 - 11.5.2 建模并实现预测 438
 - 11.5.3 效果评估 439
- 第12章 股票价格预测 442
 - 12.1 股票市场简介 443
 - 12.1.1 股票的基本概念 443
 - 12.1.2 股票市场常用术语 443
 - 12.1.3 股价波动的影响因素 447
 - 12.2 获取股票数据 452
 - 12.3 基于VAR算法的预测 455
 - 12.3.1 平稳性检验 455
 - 12.3.2 VAR模型定阶 456
 - 12.3.3 预测及效果验证 457
- 参考文献 459
- • • • • (收起)

[R语言预测实战 下载链接1](#)

标签

R语言

数据挖掘

R

计算机

统计

数据分析

大数据

评论

amazon上面的评价非常高

非常全面

图书馆借的，偏算法，数学推导写的很奇怪，推荐喜欢手动早轮子的看
堪比R语言界的《机器学习实战》

实战还是很有用的 但理论推导讲的不是很具体pa

相当不错的入门书，对数据的预测有一个不错的认识，而且代码很全，

[R语言预测实战_下载链接1](#)

书评

关键章节没有数据，无法做实验 找作者要了几次，都不回复
关键章节没有数据，无法做实验 找作者要了几次，都不回复
关键章节没有数据，无法做实验 找作者要了几次，都不回复
关键章节没有数据，无法做实验 找作者要了几次，都不回复
关键章节没有数据，无法做实验 找作者要了几...
