

<<机器学习基础教程>>

图书基本信息

书名：<<机器学习基础教程>>

13位ISBN编号：9787111407027

10位ISBN编号：7111407024

出版时间：2014-1

出版时间：机械工业出版社

作者：（英）Simon Rogers,,Mark Girolami

译者：郭茂祖,王春宇 刘扬 刘晓燕,刘扬,刘晓燕

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<机器学习基础教程>>

内容概要

本书是一本机器学习入门教程，包含了数学和统计学的核心技术，用于帮助理解一些常用的机器学习算法。

书中展示的算法涵盖了机器学习的各个重要领域：分类、聚类和投影。

本书对一小部分算法进行了详细描述和推导，而不是简单地将大量算法罗列出来。

本书通过大量的MATLAB/Octave脚本将算法和概念由抽象的等式转化为解决实际问题的工具，利用它们读者可以重新绘制书中的插图，并研究如何改变模型说明和参数取值。

本书特色

介绍机器学习技术及应用的主要算法和思想。

为读者进一步探索机器学习领域中的特定方向提供起点。

不需要太多的数学知识，穿插在文中的注解框提供相应的数学解释。

每章末均包含练习。

<<机器学习基础教程>>

作者简介

Simon Rogers 英国格拉斯哥大学计算机科学学院讲师，主讲硕士生的机器学习课程。Rogers博士是机器学习领域的一位活跃研究者，研究兴趣包括代谢组学数据分析和概率机器学习技术在人机交互领域的应用。

Mark Girolami 英国伦敦大学学院（UCL）统计系主任和计算机科学系荣誉教授，并担任计算统计学和机器学习研究中心主任。他还是英国统计协会研究组成员，英国工程和科学研究委员会高级研究员，英国工程技术学会会员，爱丁堡皇家学会院士。

<<机器学习基础教程>>

书籍目录

出版者的话

译者序

前言

第1章 线性建模：最小二乘法1

1.1 线性建模1

1.1.1 定义模型2

1.1.2 模型假设2

1.1.3 定义什么是好的模型3

1.1.4 最小二乘解：一个有效的例子4

1.1.5 有效的例子7

1.1.6 奥运会数据的最小二乘拟合8

1.1.7 小结9

1.2 预测9

1.2.1 第二个奥运会数据集10

1.2.2 小结12

1.3 向量/矩阵符号12

1.3.1 例子17

1.3.2 数值的例子18

1.3.3 预测19

1.3.4 小结19

1.4 线性模型的非线性响应19

1.5 泛化与过拟合22

1.5.1 验证数据22

1.5.2 交叉验证23

1.5.3 K折交叉验证的计算缩放25

1.6 正则化最小二乘法25

1.7 练习27

其他阅读材料28

第2章 线性建模：最大似然方法29

2.1 误差作为噪声29

2.2 随机变量和概率30

2.2.1 随机变量30

2.2.2 概率和概率分布31

2.2.3 概率的加法32

2.2.4 条件概率32

2.2.5 联合概率33

2.2.6 边缘化34

2.2.7 贝叶斯规则介绍36

2.2.8 期望值37

2.3 常见的离散分布39

2.3.1 伯努利分布39

2.3.2 二项分布39

2.3.3 多项分布40

2.4 连续型随机变量——概率密度函数40

2.5 常见的连续概率密度函数42

<<机器学习基础教程>>

- 2.5.1 均匀密度函数42
- 2.5.2 密度函数43
- 2.5.3 高斯密度函数44
- 2.5.4 多元高斯44
- 2.5.5 小结46
- 2.6 产生式的考虑(续) 46
- 2.7 似然估计47
 - 2.7.1 数据集的似然值48
 - 2.7.2 最大似然49
 - 2.7.3 最大似然解的特点50
 - 2.7.4 最大似然法适用于复杂模型52
- 2.8 偏差方差平衡问题53
- 2.9 噪声对参数估计的影响53
 - 2.9.1 参数估计的不确定性54
 - 2.9.2 与实验数据比较57
 - 2.9.3 模型参数的变异性——奥运会数据58
- 2.10 预测值的变异性59
 - 2.10.1 预测值的变异性——一个例子59
 - 2.10.2 估计值的期望值61
 - 2.10.3 小结63
- 2.11 练习63
- 其他阅读材料64
- 第3章 机器学习的贝叶斯方法66
 - 3.1 硬币游戏66
 - 3.1.1 计算正面朝上的次数67
 - 3.1.2 贝叶斯方法67
 - 3.2 精确的后验70
 - 3.3 三个场景71
 - 3.3.1 没有先验知识71
 - 3.3.2 公平的投币76
 - 3.3.3 有偏的投币78
 - 3.3.4 三个场景——总结80
 - 3.3.5 增加更多的数据80
 - 3.4 边缘似然估计80
 - 3.5 超参数82
 - 3.6 图模型83
 - 3.7 奥运会100米数据的贝叶斯处理实例84
 - 3.7.1 模型84
 - 3.7.2 似然估计85
 - 3.7.3 先验概率85
 - 3.7.4 后验概率85
 - 3.7.5 1阶多项式87
 - 3.7.6 预测89
 - 3.8 边缘似然估计用于多项式模型阶的选择90
 - 3.9 小结91
 - 3.10 练习91
 - 其他阅读材料92

<<机器学习基础教程>>

第4章 贝叶斯推理94

- 4.1 非共轭模型94
- 4.2 二值响应94
- 4.3 点估计：最大后验估计方案96
- 4.4 拉普拉斯近似100
 - 4.4.1 拉普拉斯近似实例：近似 密度101
 - 4.4.2 二值响应模型的拉普拉斯近似102
- 4.5 抽样技术103
 - 4.5.1 玩飞镖游戏104
 - 4.5.2 Metropolis-Hastings算法105
 - 4.5.3 抽样的艺术110
- 4.6 小结111
- 4.7 练习111

其他阅读材料111

第5章 分类113

- 5.1 一般问题113
- 5.2 概率分类器113
 - 5.2.1 贝叶斯分类器114
 - 5.2.2 逻辑回归121
- 5.3 非概率分类器123
 - 5.3.1 K近邻算法123
 - 5.3.2 支持向量机和其他核方法...125
 - 5.3.3 小结132
- 5.4 评价分类器的性能133
 - 5.4.1 准确率——0/1损失133
 - 5.4.2 敏感性和特异性133
 - 5.4.3 ROC曲线下的区域134
 - 5.4.4 混淆矩阵135
- 5.5 判别式和产生式分类器136
- 5.6 小结136
- 5.7 练习136

其他阅读材料137

第6章 聚类分析138

- 6.1 一般问题138
- 6.2 K均值聚类139
 - 6.2.1 聚类数目的选择141
 - 6.2.2 K均值的不足之处141
 - 6.2.3 核化K均值141
 - 6.2.4 小结144
- 6.3 混合模型144
 - 6.3.1 生成过程144
 - 6.3.2 混合模型似然函数146
 - 6.3.3 EM算法146
 - 6.3.4 例子151
 - 6.3.5 EM寻找局部最优153
 - 6.3.6 组分数目的选择153
 - 6.3.7 混合组分的其他形式154

<<机器学习基础教程>>

- 6.3.8 用EM估计MAP156
- 6.3.9 贝叶斯混合模型157
- 6.4 小结157
- 6.5 练习157
- 其他阅读材料158
- 第7章 主成分分析与隐变量模型159
 - 7.1 一般问题159
 - 7.2 主成分分析161
 - 7.2.1 选择D164
 - 7.2.2 PCA的局限性165
 - 7.3 隐变量模型165
 - 7.3.1 隐变量模型中的混合模型165
 - 7.3.2 小结166
 - 7.4 变分贝叶斯166
 - 7.4.1 选择Q () 167
 - 7.4.2 优化边界168
 - 7.5 PCA的概率模型168
 - 7.5.1 Q () 169
 - 7.5.2 $Q_{xn} (x_n)$ 170
 - 7.5.3 $Q_{wn} (w_m)$ 171
 - 7.5.4 期望值要求171
 - 7.5.5 算法172
 - 7.5.6 例子173
 - 7.6 缺失值174
 - 7.6.1 缺失值作为隐变量176
 - 7.6.2 预测缺失值176
 - 7.7 非实值数据177
 - 7.7.1 概率PPCA177
 - 7.7.2 议会数据可视化180
 - 7.8 小结184
 - 7.9 练习184
 - 其他阅读材料184
 - 词汇表185
 - 索引188

<<机器学习基础教程>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>