

# 目录

<b>第一章 机器学习概述</b>	<b>1</b>
1.1 引言 . . . . .	1
1.1.1 问题导向框架 . . . . .	1
1.1.2 数据挖掘和机器学习简介 . . . . .	3
1.2 机器学习的分类 . . . . .	4
1.2.1 无监督学习和有监督学习 . . . . .	4
1.2.2 强化学习 . . . . .	5
1.2.3 浅层学习和深度学习 . . . . .	6
1.2.4 人工智能简介 . . . . .	6
1.3 统计学习 . . . . .	7
<b>第二章 Python 科学计算简介</b>	<b>9</b>
2.1 基础变量类型 . . . . .	9
2.1.1 Number 数字 . . . . .	10
2.1.2 String 字符串 . . . . .	11
2.1.3 List 列表 . . . . .	12
2.1.4 Tuple 元组 . . . . .	14
2.1.5 Dictionary 字典 . . . . .	14
2.2 控制语句和函数 . . . . .	15
2.2.1 控制语句 . . . . .	15
2.2.2 函数 . . . . .	18
2.3 科学计算和数据处理 . . . . .	19
2.3.1 Numpy . . . . .	20

目录	4
2.3.2 SciPy	24
2.3.3 Pandas	25
2.4 作图和可视化	27
2.4.1 Plot函数	27
2.4.2 标题、图例和坐标	28
2.4.3 常用绘图函数	28
2.4.4 Image Plot	29
2.5 输入和输出	30
2.5.1 标准输入和输出	30
2.5.2 第三方库的输入输出函数	31
2.5.3 案例分析：读取并处理股票行情数据	31
2.6 面向对象编程	32
2.6.1 过程式编程	33
2.6.2 案例分析：面向对象编程示例	33
2.7 常用工具库和部署	34
习题	36
<b>第三章 无监督学习基础</b>	<b>38</b>
3.1 描述性统计	38
3.1.1 描述性统计分析工具	38
3.1.2 案例分析：指数收益率的描述性统计	41
3.2 核密度估计	42
3.2.1 核密度估计方法	42
3.2.2 核密度估计的目标函数	44
3.3 K均值算法	44
3.4 主成分分析	47
3.4.1 最大投影方差和最小重构误差	47
3.4.2 特征分解和奇异值分解	49
3.4.3 案例分析：手写数字3特征分析	50
3.4.4 案例分析：利率期限结构	53

目录	5
3.4.5 案例分析：股票收益率的协方差矩阵分解	55
3.5 混合模型和隐马尔科夫模型	57
3.5.1 混合模型	57
3.5.2 隐马尔科夫模型	58
习题	62
<b>第四章 线性回归和正则化方法</b>	<b>64</b>
4.1 回归分析流程	64
4.1.1 回归分析流程主要步骤	65
4.1.2 案例分析：宏观违约率预测	69
4.2 变量选择基础	71
4.2.1 子集选择	71
4.2.2 案例分析：指数跟踪	72
4.2.3 Forward Stagewise 回归	73
4.3 正则化方法	75
4.3.1 L2正则	76
4.3.2 L1正则	77
4.3.3 惩罚函数和稀疏性	84
4.4 回归估计和矩阵分解	86
4.4.1 奇异值分解和线性回归	86
4.4.2 QR分解和QR算法	89
习题	91
<b>第五章 分类</b>	<b>93</b>
5.1 判别分析	93
5.1.1 线性判别分析	94
5.1.2 二次判别分析	95
5.1.3 朴素贝叶斯	96
5.2 逻辑回归	96
5.2.1 模型估计	97

5.2.2	与交叉熵的关系 . . . . .	101
5.2.3	案例分析：股票涨跌预测 . . . . .	102
5.3	支持向量机 . . . . .	104
5.4	分类的评判 . . . . .	106
5.4.1	混淆矩阵和常用度量 . . . . .	107
5.4.2	F1 Score . . . . .	109
5.4.3	ROC 和AUC . . . . .	109
5.4.4	数据不平衡的处理 . . . . .	112
	习题 . . . . .	113
<b>第六章</b>	<b>局部建模</b>	<b>115</b>
6.1	样条方法 . . . . .	115
6.1.1	三阶样条 . . . . .	115
6.1.2	自然三阶样条 . . . . .	117
6.2	核技巧 . . . . .	118
6.3	局部回归 . . . . .	120
6.3.1	K邻近估计 . . . . .	120
6.3.2	局部常数估计 . . . . .	122
6.3.3	局部多项式估计 . . . . .	123
6.3.4	案例分析：期权隐含分布估计 . . . . .	125
6.3.5	局部似然估计 . . . . .	127
	习题 . . . . .	128
<b>第七章</b>	<b>模型选择和模型评估</b>	<b>129</b>
7.1	模型评估 . . . . .	129
7.1.1	泛化误差 . . . . .	129
7.1.2	交叉验证 . . . . .	131
7.1.3	Bootstrap . . . . .	133
7.2	模型选择 . . . . .	133
7.2.1	AIC 准则 . . . . .	134

7.2.2	BIC 准则	136
7.3	估计的自由度	138
7.4	案例分析：期权隐含分布估计（续1）	139
	习题	140
<b>第八章</b>	<b>统计推断基础</b>	<b>142</b>
8.1	极大似然估计	142
8.2	置信区间和假设检验	144
8.2.1	置信区间	144
8.2.2	假设检验	144
8.3	Bootstrap方法	146
8.4	KL距离和信息论相关概念	150
8.4.1	KL距离和熵	150
8.4.2	KL距离和互信息	151
8.5	EM算法	152
8.5.1	EM算法与变分推断和MM算法	153
8.5.2	高斯混合模型的EM算法	154
8.5.3	隐马尔科夫模型的EM算法	156
8.5.4	案例分析：收益率序列隐状态预测	161
	习题	163
<b>第九章</b>	<b>贝叶斯方法</b>	<b>164</b>
9.1	贝叶斯定理	164
9.1.1	事件的贝叶斯公式	164
9.1.2	随机变量的贝叶斯公式	165
9.2	贝叶斯视角下的频率方法	166
9.3	抽样方法	168
9.3.1	拒绝抽样法	168
9.3.2	案例分析：期权隐含分布估计（续2）	169
9.3.3	Metropolis-Hastings 抽样算法	170

9.3.4 重要性抽样 . . . . .	176
9.3.5 蒙特卡洛标准误 . . . . .	178
9.4 变分推断 . . . . .	178
9.4.1 基于平均场的变分推断 . . . . .	178
9.4.2 变分推断算法示例 . . . . .	180
习题 . . . . .	184
<b>第十章 树和树的集成</b>	<b>186</b>
10.1 回归树和分类树 . . . . .	186
10.1.1 回归树 . . . . .	186
10.1.2 分类树 . . . . .	189
10.2 Bagging和随机森林 . . . . .	193
10.2.1 Bagging . . . . .	193
10.2.2 随机森林 . . . . .	194
10.3 提升树Boosting Trees . . . . .	196
10.3.1 AdaBoost . . . . .	196
10.3.2 梯度提升树GBDT . . . . .	197
10.3.3 XGBoost . . . . .	199
10.3.4 案例分析：股票涨跌预测（续1） . . . . .	201
习题 . . . . .	203
<b>第十一章 深度学习</b>	<b>204</b>
11.1 前馈神经网络和梯度下降算法 . . . . .	204
11.1.1 神经元 . . . . .	204
11.1.2 前馈神经网络 . . . . .	206
11.1.3 梯度下降算法 . . . . .	207
11.1.4 反向传播算法 . . . . .	207
11.1.5 随机梯度算法的改进 . . . . .	209
11.1.6 激活函数和梯度消失问题 . . . . .	210
11.1.7 案例分析：股票涨跌预测（续2） . . . . .	212

11.2 网络结构 . . . . .	214
11.2.1 卷积神经网络CNN . . . . .	214
11.2.2 循环神经网络RNN . . . . .	218
11.2.3 Dropout . . . . .	220
11.2.4 Batch Normalization . . . . .	220
11.2.5 残差网络 . . . . .	221
11.3 自编码和生成模型 . . . . .	222
11.3.1 自编码 . . . . .	222
11.3.2 案例分析：手写数字3特征分析（续） . . . . .	224
11.3.3 逐层特征学习 . . . . .	225
11.3.4 生成对抗网络 . . . . .	227
11.3.5 变分自编码 . . . . .	227
11.4 揭开深度学习的黑箱？ . . . . .	229
习题 . . . . .	232
<b>第十二章 深度强化学习</b>	<b>233</b>
12.1 基于值函数的强化学习 . . . . .	233
12.1.1 强化学习的基础概念 . . . . .	233
12.1.2 值函数和Bellman方程 . . . . .	234
12.1.3 策略迭代和值迭代 . . . . .	236
12.1.4 基于值函数的无模型强化学习 . . . . .	237
12.2 值函数近似和深度Q网络 . . . . .	241
12.2.1 值函数的近似 . . . . .	241
12.2.2 深度Q网络DQN . . . . .	242
12.2.3 案例分析：DQN智能交易机器人 . . . . .	244
12.3 策略梯度和Actor-Critic方法 . . . . .	246
12.3.1 策略梯度定理 . . . . .	246
12.3.2 强化学习和有监督学习的对比 . . . . .	248
12.3.3 Actor-Critic算法 . . . . .	249
12.4 学习、推演和搜索 . . . . .	252

12.4.1 “记忆式”学习 . . . . .	252
12.4.2 推演和搜索 . . . . .	252
12.4.3 蒙特卡洛树搜索 . . . . .	253
12.4.4 不完全信息决策简介 . . . . .	254
习题 . . . . .	255