

MACHINE
LEARNING WITH PYTHON

Python 机器学习

[印] 阿布舍克·维贾亚瓦吉亚 (Abhishek Vijayvargia) 著

宋格格 译

- 走进机器学习
- 了解 Python
- 特征工程
- 数据可视化
- 基本和高级回归技术
- 分类
- 无监督学习
- 文本分析
- 神经网络和深度学习
- 推荐系统
- 时间序列分析

199.

MACHINE
LEARNING WITH PYTHON

Python 机器学习

[印] 阿布舍克·维贾亚瓦吉亚 (Abhishek Vijayvargia) 著

宋格格 译

人民邮电出版社
北京

图书在版编目(CIP)数据

Python机器学习 / (印) 阿布舍克·维贾亚瓦吉亚
(Abhishek Vijayvargia) 著; 宋格格译. — 北京: 人
民邮电出版社, 2019.2
ISBN 978-7-115-50135-6

I. ①P… II. ①阿… ②宋… III. ①软件工具—程序
设计 IV. ①TP311.561

中国版本图书馆CIP数据核字(2018)第264219号

版权声明

Original edition published by BPB Publications. Copyright © 2018 by BPB Publications. Simplified Chinese-language edition copyright © 2018 by POSTS & TELECOM PRESS. All rights reserved.

本书中文简体字版由印度 **BPB Publications** 授权人民邮电出版社出版。未经出版者书面许可, 对本书任何部分不得以任何方式复制或抄袭。

版权所有, 侵权必究。

-
- ◆ 著 [印] 阿布舍克·维贾亚瓦吉亚 (Abhishek Vijayvargia)
译 宋格格
责任编辑 陈聪聪
责任印制 焦志炜
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
涿州市京南印刷厂印刷
 - ◆ 开本: 800×1000 1/16
印张: 17.75 彩插: 4
字数: 314千字 2019年2月第1版
印数: 1-2400册 2019年2月河北第1次印刷
- 著作权合同登记号 图字: 01-2018-4177 号
-

定价: 79.00元

读者服务热线: (010)81055410 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147号

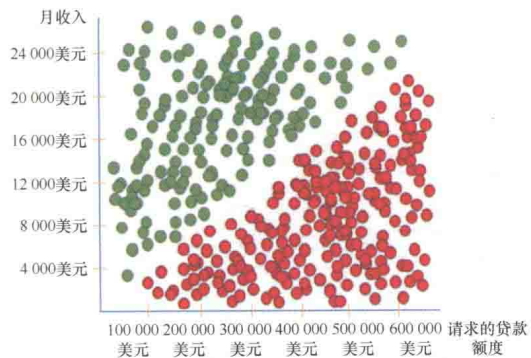


图1.18

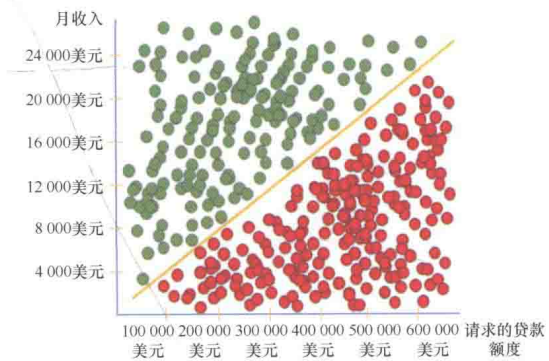


图1.19

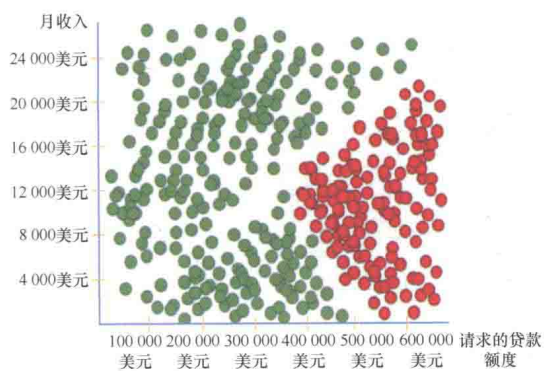


图1.20

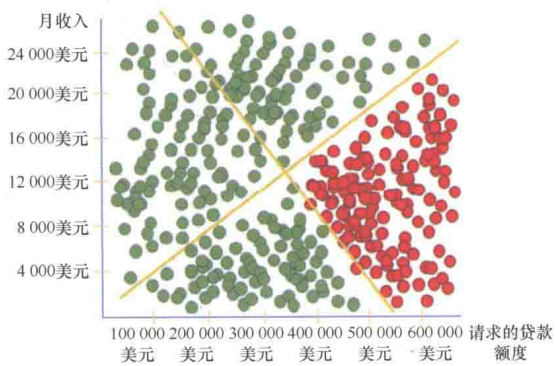


图1.21

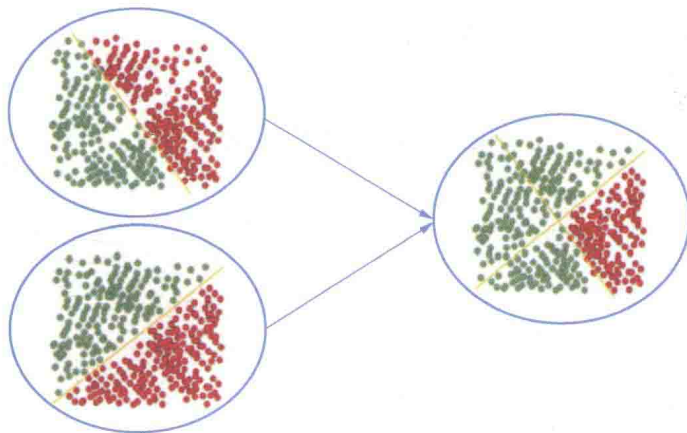


图1.22



图1.25

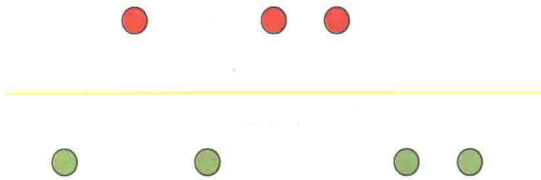


图1.26

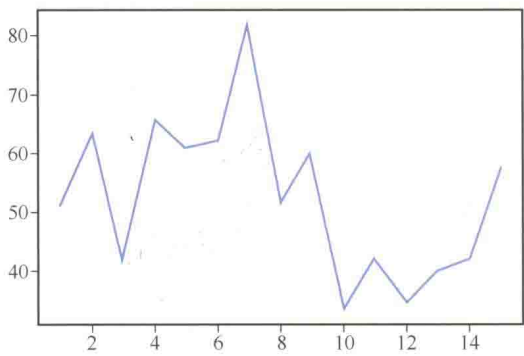


图4.1

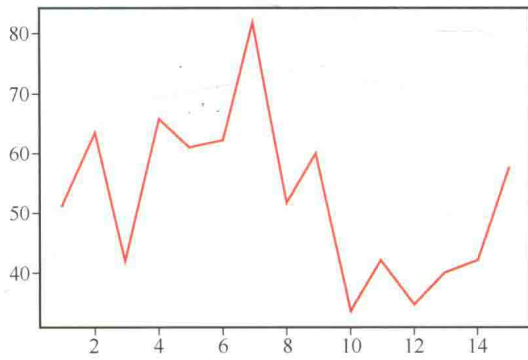


图4.2

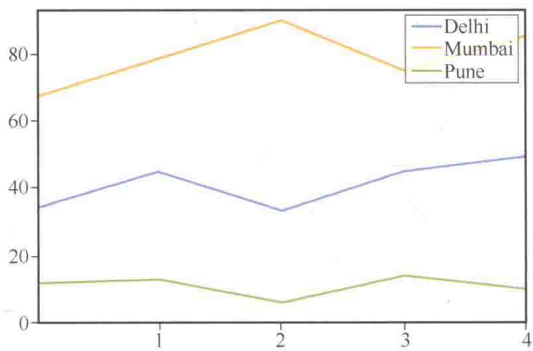


图4.5

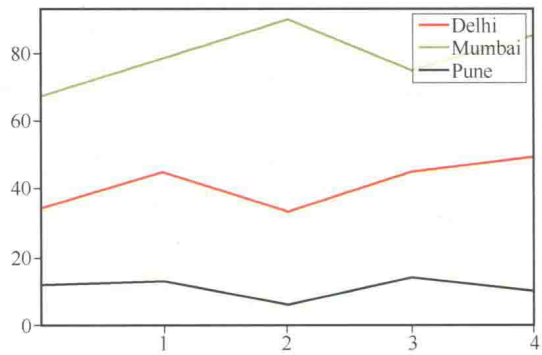


图4.6

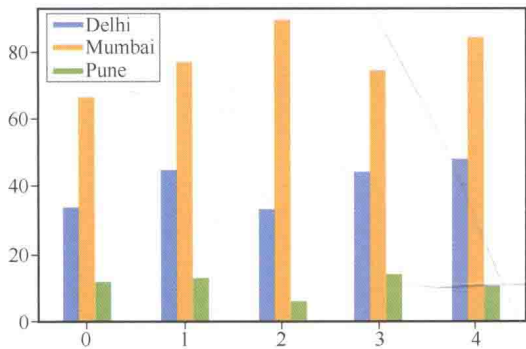


图4.8

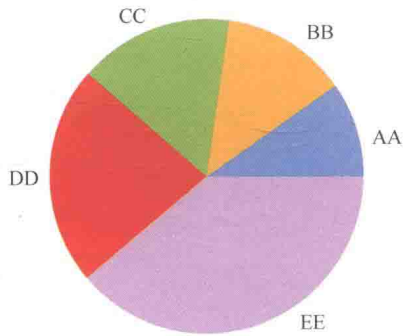


图4.9

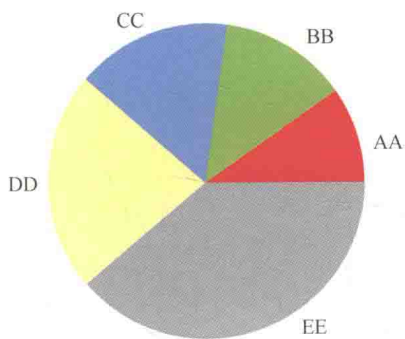


图 4.10

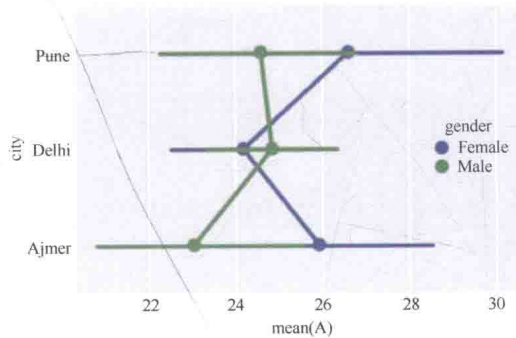


图 4.22

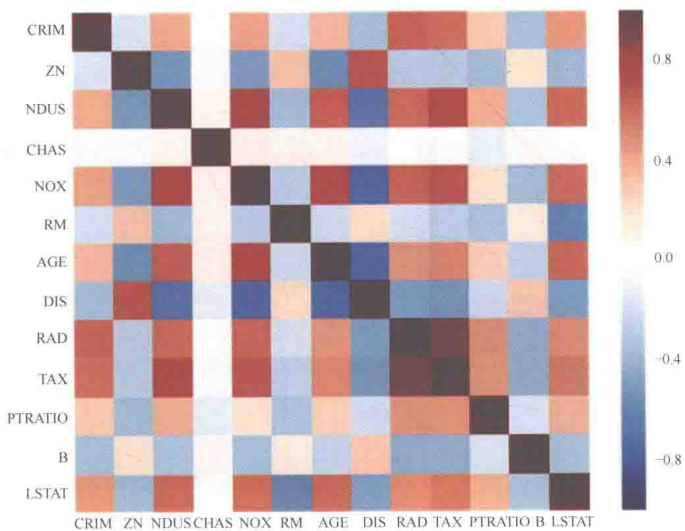


图 5.21

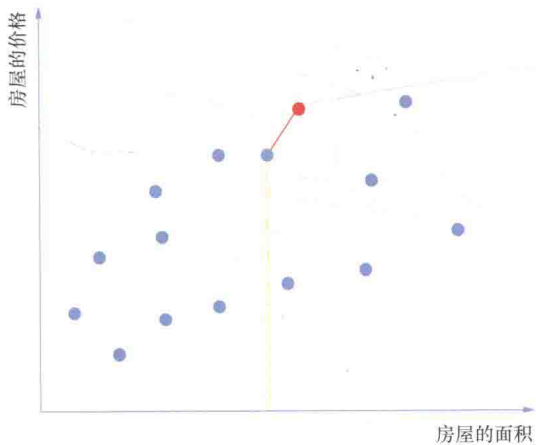


图 6.16

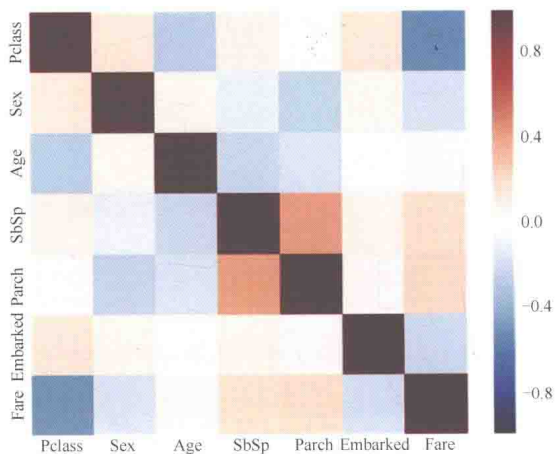


图 7.11

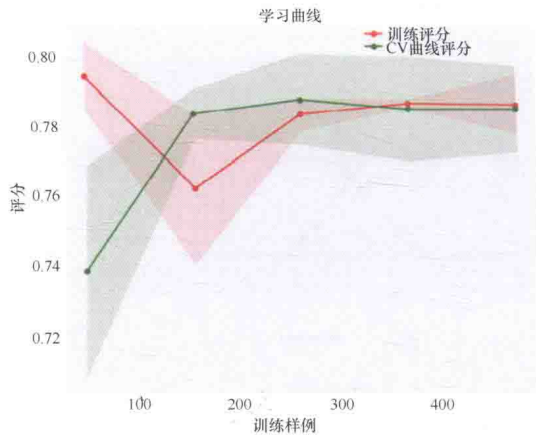


图 7.12

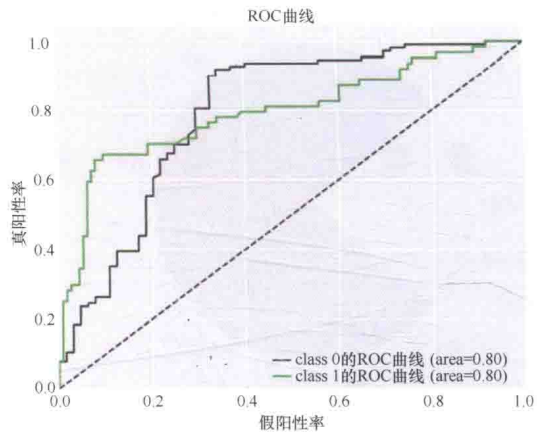


图 7.13

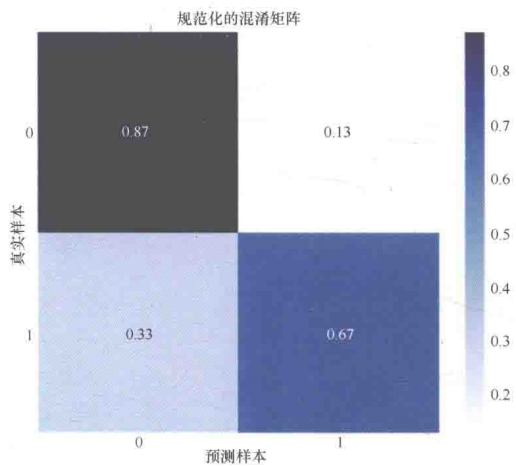


图 7.14

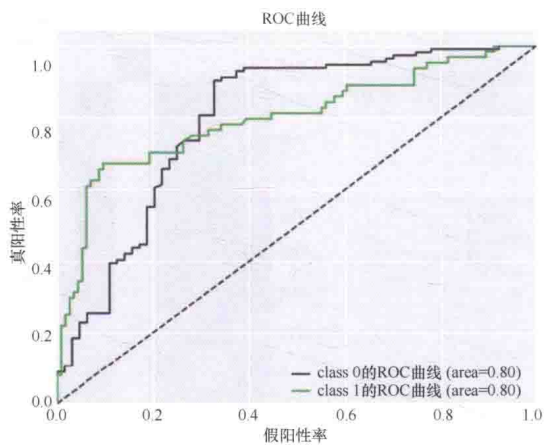


图 7.15

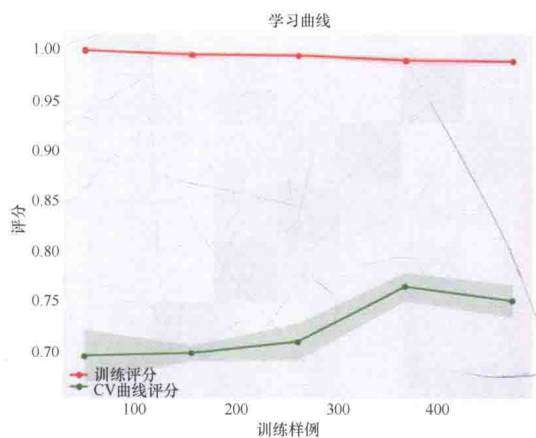


图 7.22

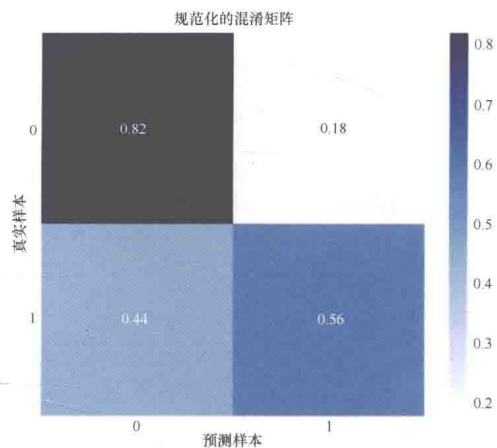


图 7.23

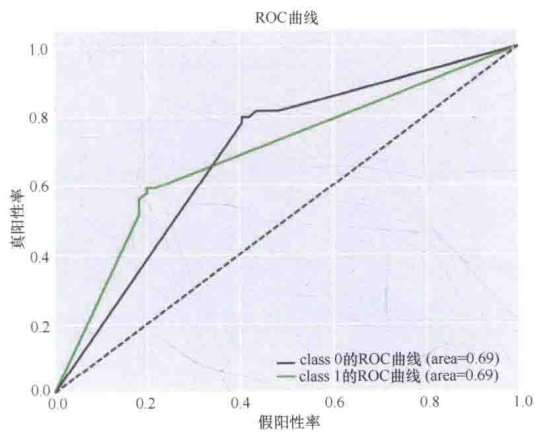


图 7.24

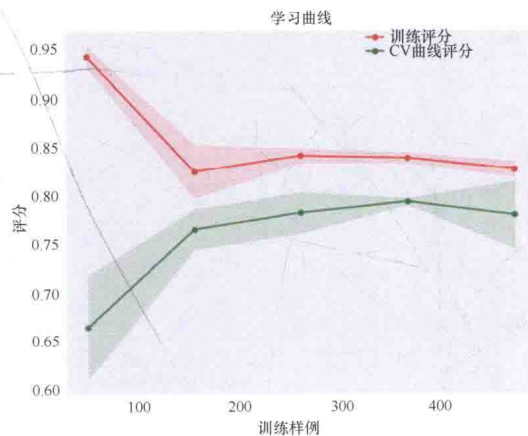


图 7.25

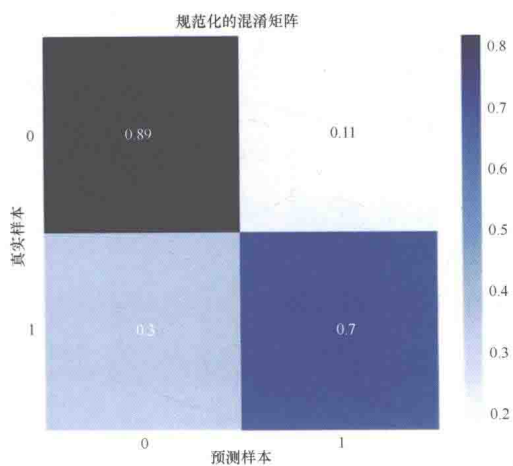


图 7.26

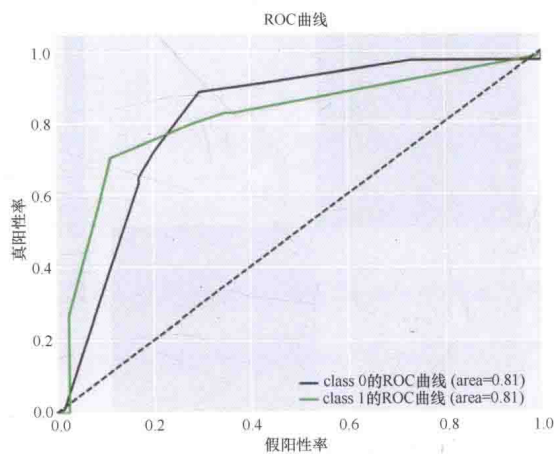


图 7.27

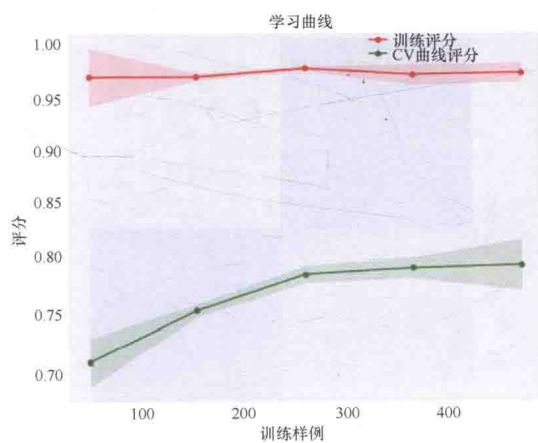


图 7.29

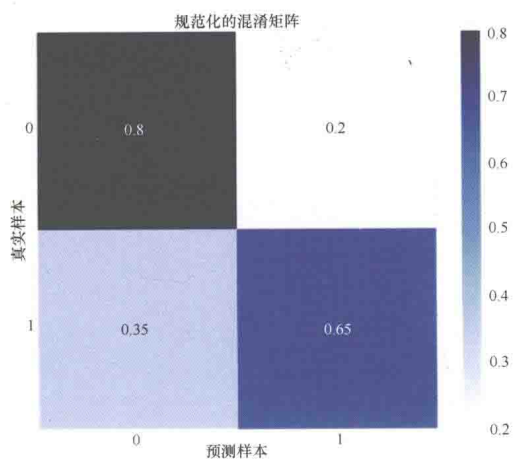


图 7.30

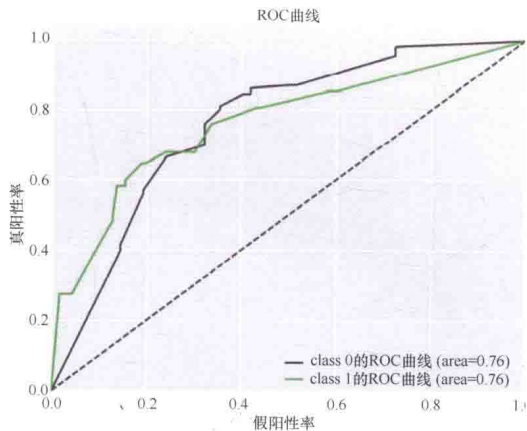


图 7.31

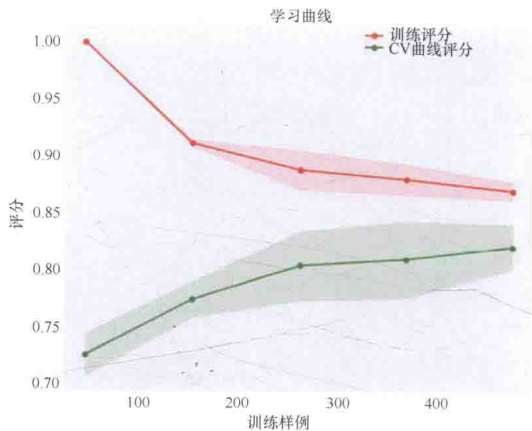


图 7.32

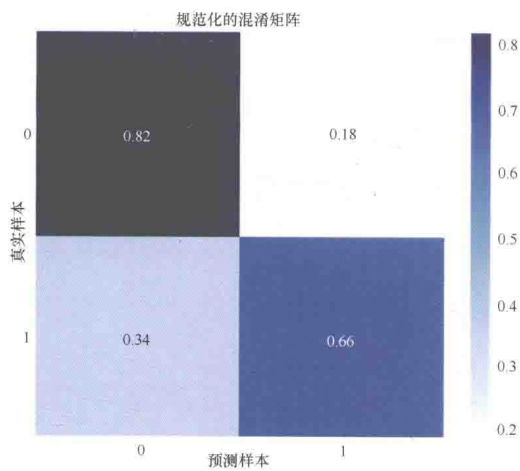


图 7.33

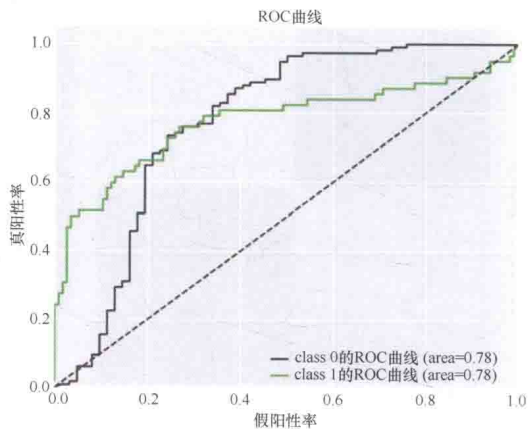


图 7.34

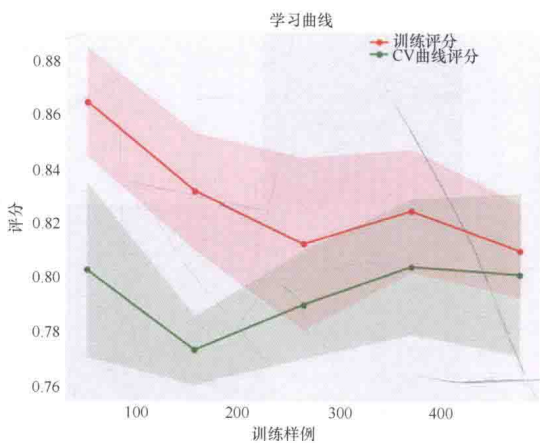


图 7.35

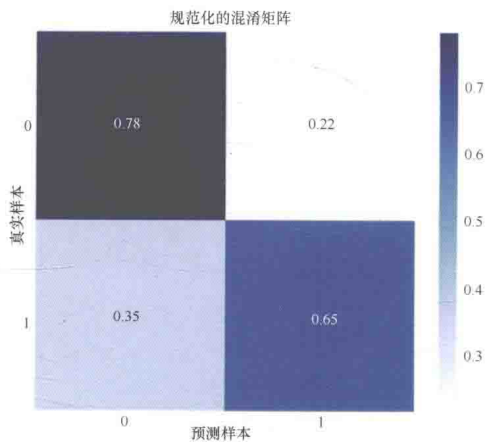


图 7.36

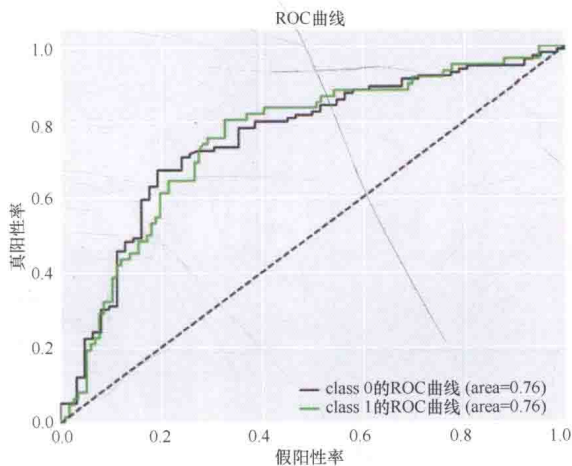


图 7.37

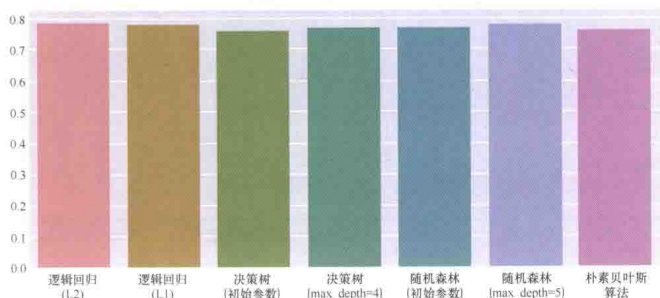


图 7.38

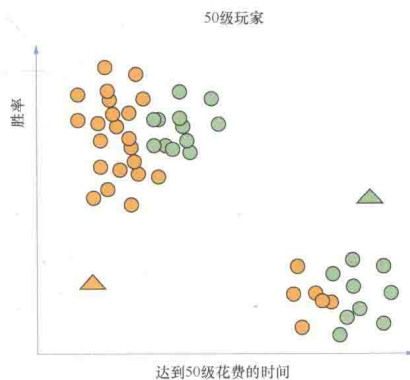


图 8.2

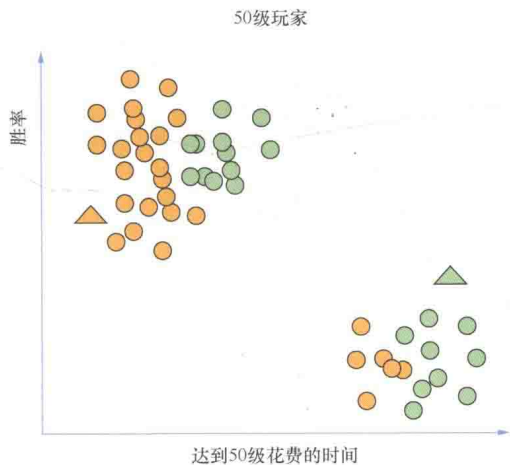


图 8.3

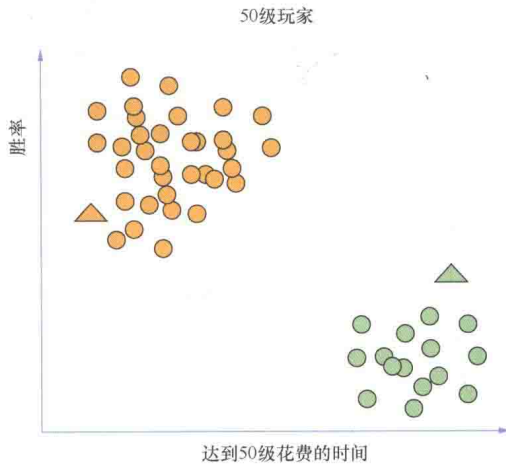


图 8.4

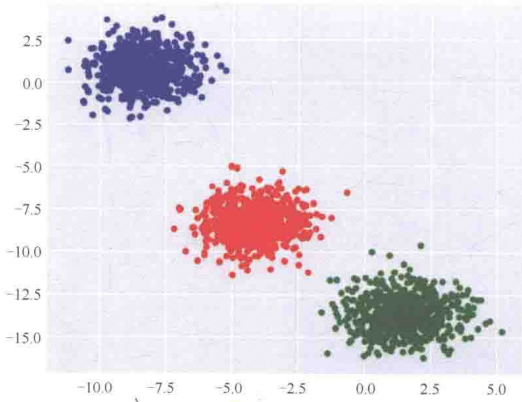


图 8.10

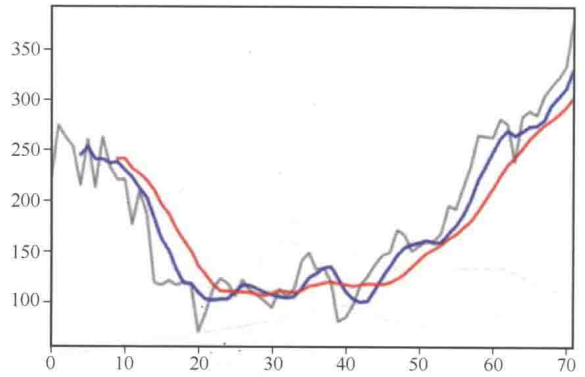


图 12.2

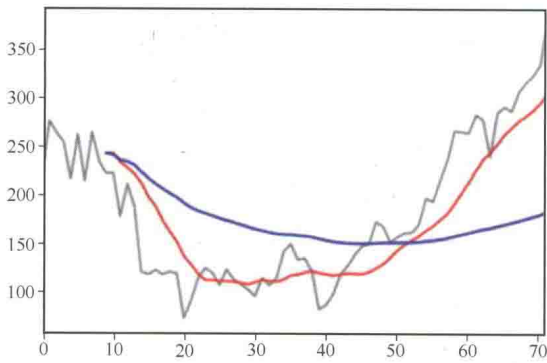


图 12.7

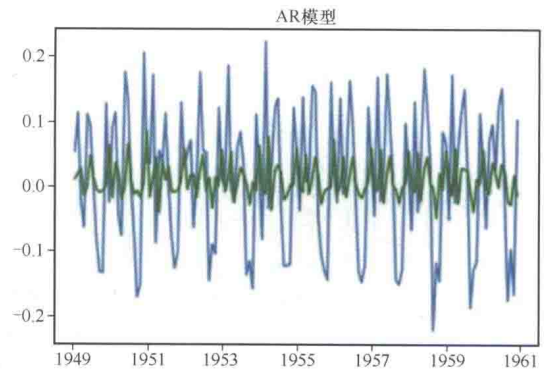


图 12.16

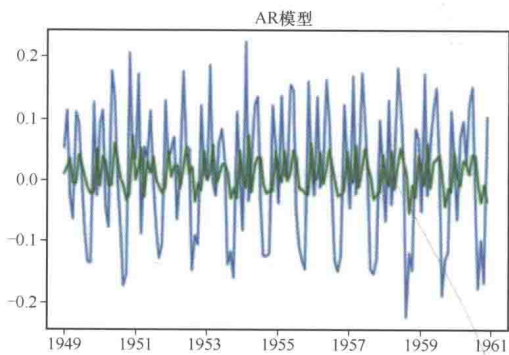


图 12.17

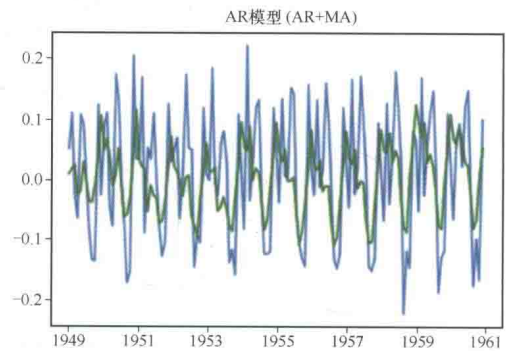


图 12.18

内容提要

本书通过解释数学原理和展示编程示例对机器学习进行了系统、全面的解析。全书共分为12章，内容涵盖了机器学习以及Python语言的基础知识、特征工程的概念与操作技术、数据可视化技术的实现、监督学习及无监督学习算法、文本分析、神经网络和深度学习、推荐系统的构建方法以及预测处理时间序列的方法等。阅读本书能够加深读者对机器学习的认识和理解，从而达到理论与实践相结合、学以致用目的。

本书适合Python程序员、数据分析人员、对机器学习感兴趣的读者以及机器学习领域的从业人员阅读。

感谢

感谢BPB出版社使这本书得以出版。

非常感谢我的家人和朋友，没有他们的支持，我将无法完成这项任务。

“无论我走多远，我的父母都会永远和我在一起。”

——Abhishek Vijayvargia

前言

机器学习是计算机科学领域未来发展的重点之一。从制造业到电子商务、再到运输业、机器学习无处不在，并提供基于数据的解决方案。机器学习背后的理念是在决策过程中建立一个与人类一样聪明的知识模型。现在，随着计算和存储能力的增强，机器学习相比人类可以提供更好的解决方案。

本书由12章组成，每一章涉及机器学习的一个主要领域。

第1章从机器学习的基本介绍开始，读者可以结合实际应用案例来了解通用机器学习算法的思想。第2章介绍了Python语言的基础知识，Python是一种开源语言，也是一种很好的机器学习应用工具，这是选择它来编写机器学习算法的原因。第3章提供了特征工程的相关操作技术，对机器学习而言，寻找正确的特征并加以修改与算法同样重要。第4章重点介绍数据可视化技术，使用预先构建的Python库，用户可以将数据可视化并向他人展示。

第5章~第7章涵盖了监督学习算法的内容。第5章通过实例解释基本回归技术。第6章通过拟合问题的解决方案来重点介绍高级回归技术。第7章描述了分类算法的细节，从参数化技术和非参数化技术两方面来探讨解决分类问题。第8章引入无监督学习的思想，重点介绍了聚类。

第9章~第12章提供了机器学习中一些高级概念。第9章介绍了文本分析，包含一个示例——将新闻分类到预定义类别中。第10章讨论了神经网络和深度学习，其在非结构化、图像和语音数据中得到了高度的应用。第11章介绍了构建推荐系统的方法，对每种方法给出了相应的案例。第12章讨论时间序列数据、处理和预测时间序列的方法。

本书对各学习阶段的读者都有帮助。无论读者是刚开始接触机器学习，还是想学习更深层的概念，又或者是想通过代码进行实践，本书均可满足需求。建议读者通过使用示例

代码来学习概念并实践它，以掌握这本书的全部内容。

虽然其他数据科学家已经对书中的内容和代码进行了检查，但本书仍然可能存在一些缺陷。期待读者的建议和批评，我将尽力改正以求在未来的版本中不再出现这些错误。

资源与支持

本书由异步社区出品，社区 (<https://www.epubit.com/>) 为您提供相关资源和后续服务。

配套资源

本书提供如下资源：

- 本书配套资源请到异步社区本书购买页处下载。

要获得以上配套资源，请在异步社区本书页面中点击 **配套资源**，跳转到下载界面，按提示进行操作即可。注意：为保证购书读者的权益，该操作会给出相关提示，要求输入提取码进行验证。

提交勘误

作者和编辑尽最大努力来确保书中内容的准确性，但难免会存在疏漏。欢迎您将发现的问题反馈给我们，帮助我们提升图书的质量。

当您发现错误时，请登录异步社区，按书名搜索，进入本书页面，点击“提交勘误”，输入勘误信息，单击“提交”按钮即可。本书的作者和编辑会对您提交的勘误进行审核，确认并接受后，您将获赠异步社区的100积分。积分可用于在异步社区兑换优惠券、样书或奖品。



扫码关注本书

扫描下方二维码，您将会在异步社区微信服务号中看到本书信息及相关的服务提示。



与我们联系

我们的联系邮箱是 contact@epubit.com.cn。

如果您对本书有任何疑问或建议，请您发邮件给我们，并请在邮件标题中注明本书书名，以便我们更高效地做出反馈。

如果您有兴趣出版图书、录制教学视频，或者参与图书翻译、技术审校等工作，可以发邮件给我们；有意出版图书的作者也可以到异步社区在线提交投稿（直接访问 www.epubit.com/selfpublish/submission 即可）。

如果您是学校、培训机构或企业，想批量购买本书或异步社区出版的其他图书，也可以发邮件给我们。

如果您在网上发现有针对异步社区出品图书的各种形式的盗版行为，包括对图书全部或部分内容的非授权传播，请您将怀疑有侵权行为的链接发邮件给我们。您的这一举动是对作者权益的保护，也是我们持续为您提供有价值的内容的动力之源。

关于异步社区和异步图书

“异步社区”是人民邮电出版社旗下IT专业图书社区，致力于出版精品IT技术图书和相关学习产品，为作译者提供优质出版服务。异步社区创办于2015年8月，提供大量精品IT技术图书和电子书，以及高品质技术文章和视频课程。更多详情请访问异步社区官网 <https://www.epubit.com>。

“异步图书”是由异步社区编辑团队策划出版的精品IT专业图书的品牌，依托于人民邮电出版社近30年的计算机图书出版积累和专业编辑团队，相关图书在封面上印有异步图书的LOGO。异步图书的出版领域包括软件开发、大数据、AI、测试、前端、网络技术 etc。



异步社区



微信服务号