

PACKT
PUBLISHING

直观的数据项目，应用高级机器学习方法解决日常问题

Python 机器学习 实践指南

Python Machine
Learning Blueprints

[美] Alexander T. Combs 著
黄申 译

 中国工信出版集团

 人民邮电出版社
POSTS & TELECOM PRESS



Python机器学习 实践指南

[美] Alexander T. Combs 著
黄申 译

人民邮电出版社

图书在版编目 (C I P) 数据

Python机器学习实践指南 / (美) 库姆斯
(Alexander T. Combs) 著 ; 黄申译. -- 北京 : 人民邮
电出版社, 2017. 5

书名原文: Python Machine Learning Blueprints
ISBN 978-7-115-44906-1

I. ①P… II. ①库… ②黄… III. ①软件工具—程序
设计—指南 IV. ①TP311.561-62

中国版本图书馆CIP数据核字 (2017) 第050069号

-
- ◆ 著 [美] Alexander T. Combs
译 黄 申
责任编辑 陈冀康
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>
北京市艺辉印刷有限公司印刷
 - ◆ 开本: 800×1000 1/16
印张: 17
字数: 330 千字 2017 年 5 月第 1 版
印数: 1 - 3 000 册 2017 年 5 月北京第 1 次印刷
- 著作权合同登记号 图字: 01-2016-7608 号
-

定价: 69.00 元

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

()

广告经营许可证: 京东工商广字第 8052 号

内容提要

机器学习是近年来渐趋热门的一个领域，同时 Python 语言经过一段时间的发展也已逐渐成为主流的编程语言之一。本书结合了机器学习和 Python 语言两个热门的领域，通过易于理解的项目详细讲述了如何构建真实的机器学习应用程序。

全书共有 10 章。第 1 章讲解了 Python 机器学习的生态系统，剩余 9 章介绍了众多与机器学习相关的算法，包括聚类算法、推荐引擎等，主要包括机器学习在公寓、机票、IPO 市场、新闻源、内容推广、股票市场、图像、聊天机器人和推荐引擎等方面的应用。

本书适合 Python 程序员、数据分析人员、对算法感兴趣的读者、机器学习领域的从业人员及科研人员阅读。

作者简介

Alexander T. Combs 是一位经验丰富的数据科学家、策略师和开发人员。他有金融数据抽取、自然语言处理和生成，以及定量和统计建模的背景。他目前是纽约沉浸式数据科学项目的一名全职资深讲师。

审阅者简介

KushalKhandelwal 是一位数据科学家和全栈开发人员。他的兴趣包括构建可扩展的机器学习和图像处理的软件应用。他擅长 Python 编码，并对各种开源项目做出了积极的贡献。他目前担任 Truce.in 的技术主管，这是一家以农民为中心的创业公司，Kushal 致力于创建可扩展的 Web 应用程序来帮助农民。

译者简介



黄申博士，现任 IBM 研究院资深科学家，毕业于上海交通大学计算机科学与工程专业，师从俞勇教授。微软学者、IBM ExtremeBlue 天才计划成员。长期专注于大数据相关的搜索、推荐、广告以及用户精准化领域。曾在微软亚洲研究院、eBay 中国、沃尔玛 1 号店（现京东 1 号店）和大润发飞牛网担任要职，带队完成了若干公司级的战略项目。同时发表了 20 多篇国际论文，并拥有 10 多项国际专利，《计算机工程》特邀审稿专家，《Elasticsearch 实战》中文版的译者，2016 年出版的《大数据架构商业之路》一书销量和口碑双赢，续作《大数据架构和算法实现之路》将于 2017 年中出版。2015 年，因对业界做出卓越贡献，获得美国政府颁发的“美国杰出人才”称号。

译者序

谈到为什么要翻译这本书，还是一段机缘巧合。那是 2015 年的下半年，当时我正在撰写自己的原创书籍《大数据架构商业之路：从业务需求到技术方案》。在那本书中，我希望结合一个创业的故事，展示各个阶段可能遇到的大数据课题、业务需求，以及相对应的技术方案，甚至是实践解析。其中，最挑战的部分莫过于案例的分析到技术方案，再到框架编码的逐步展开。因为之前对于这种写作模式没有相关的经验，让人很是苦恼。我也搜寻了市面上相关的中英文书籍，可惜并未发现特别好的范例作为参考。

一次偶然的的机会，我在 Amazon.com 上发现了 Alexander T. Combs 的《Python Machine Learning Blueprints》。当时此书尚未出版，还是试读本。在阅读样章之后我发现这种写作模式就是我要的，没有太多的理论和说教，而是结合我们日常生活都会经历的方方面面，包括房产、金融、旅游和电子商务等，提供了可以直接上手的教学内容，让读者可以身临其境，乐在其中，轻松了解机器学习的实用知识。这正是我想要学习的风格！于是我采纳了这种模式，并结合自己的项目经验，一口气完成了《大数据架构商业之路：从业务需求到技术方案》一书。上市之后，读者对这种理论和案例相结合的方式很是赞许。所以，我对《Python Machine Learning Blueprints》一书心存感激，对它何时上市也很是关注。

终于，2016 年的 7 月底，该书的英文版正式发行。我迫不及待地阅读完了原版，和当初试读的感觉一样，这是一本很有创意的书，而且 Python 和机器学习都是最近几年的技术热点，如果能将这么棒的内容介绍给广大国内的读者，那是多么令人激动的事情！于是，我抱着试试看的心态，联系了人民邮电出版社的编辑陈冀康老师。很幸运，当时此书还没有译者，陈老师审阅我的试译稿之后也表示满意，于是我很荣幸地成为了此书的译者。

不过在翻译的过程中，我也发现了不少细节上的疑问，于是我主动联系了原书的作者 Alexander，他总是非常仔细地解答这些问题，使得我信心大增，可以确保译文尽可能地贴

近原文。而编辑陈老师也对此举表示了充分的肯定。在此，我要对 Alexander 和陈老师的帮助表示衷心的感谢。当然，我也要感谢父母和妻儿的支持，为了此书，我陪伴你们的时间更少了，而你们丝毫没有怨言，让我可以安心地完成每次的写作。

在翻译此书的岁月中，Python、机器学习及其应用在国内外都获得了空前的关注，相关的社区也保持了非常好的活跃度，相信这个技术方向在将来还有很大的空间。希望本书能帮助到每一位热爱 Python 和机器学习的朋友，为中国的人工智能事业尽一份绵薄之力。如果您对本书中的技术细节感兴趣，可以通过如下渠道联系我，很期待和大家的互动和交流。

QQ 36638279

微信 18616692855

邮箱 s_huang790228@hotmail.com

LinkedIn <https://cn.linkedin.com/in/shuang790228>



扫一扫就能微信联系作者：

个人公众号

前言

机器学习正在迅速成为数据驱动型世界的一个必备模块。许多不同的领域如机器人、医学、零售和出版等，都需要依赖这门技术。在这本书中，你将学习如何一步步构建真实的机器学习应用程序。

通过易于理解的项目，你将学习如何处理各种类型的数据，如何以及何时应用不同的机器学习技术，包括监督学习和无监督学习。

本书中的每个项目都同时提供了教学和实践。例如，你将学习如何使用聚类技术来发现低价的机票，以及如何使用线性回归找到一间便宜的公寓。本书以通俗易懂、简洁明了的方式，教你如何使用机器学习来收集、分析并操作大量的数据。

本书涵盖的内容

第 1 章，Python 机器学习的生态系统，深入 Python，它有一个深度活跃的开发者社区，而且许多开发者来自科学社区。这为 Python 提供了丰富的科学计算库。在本章中，我们将讨论这些关键库的特性以及如何准备你的环境，以最好地利用它们。

第 2 章，构建应用程序，发现低价的公寓，指导我们构建第一个机器学习应用程序，我们从一个最小但实际的例子开始：建设应用程序来识别低价的公寓。到本章结束，我们将创建一个应用程序，使得寻找合适的公寓变得更容易点。

第 3 章，构建应用程序，发现低价的机票，演示了如何构建应用程序来不断地监测票价。一旦出现异常价格，应用程序将提醒我们，可以快速采取行动。

第 4 章，使用逻辑回归预测 IPO 市场，展示了我们如何使用机器学习决定哪些 IPO 值得仔细研究，而哪些可以直接跳过。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：《Python机器学习实践指南》库姆斯 (Alexander T. Combs) 著.pdf

请登录 <https://shgis.com/post/4049.html> 下载完整文档。

手机端请扫码查看：

