

O'REILLY®



# 深度学习实战

Deep Learning Cookbook: Practical Recipes to Get Started Quickly

 机械工业出版社  
China Machine Press

[美] Douwe Osinga 著  
李君婷 闫龙川 俞学豪 高德荃 译

O'Reilly 精品图书系列

---

# 深度学习实战

## Deep Learning Cookbook

[美] 杜威·奥辛格 (Douwe Osinga) 著  
李君婷 闫龙川 俞学豪 高德荃 译

Beijing · Boston · Farnham · Sebastopol · Tokyo

**O'REILLY®**

lia, Inc. 授权机械工业出版社出版

机械工业出版社

## 图书在版编目 (CIP) 数据

深度学习实战 / (美) 杜威·奥辛格 (Douwe Osinga) 著; 李君婷等译. —北京: 机械工业出版社, 2019.4 (2019.6 重印)

(O'Reilly 精品图书系列)

书名原文: Deep Learning Cookbook

ISBN 978-7-111-62483-7

I. 深… II. ①杜… ②李… III. 机器学习 IV. TP181

中国版本图书馆 CIP 数据核字 (2019) 第 068514 号

北京市版权局著作权合同登记

图字: 01-2019-1166 号

Copyright © 2018 Douwe Osinga. All rights reserved.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and China Machine Press, 2019. Authorized translation of the English edition, 2018 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly Media, Inc. 出版 2018。

简体中文版由机械工业出版社出版 2019。英文原版的翻译得到 O'Reilly Media, Inc. 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc. 的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式重制。

封底无防伪标均为盗版

本书法律顾问

北京大成律师事务所 韩光 / 邹晓东

书 名 / 深度学习实战

书 号 / ISBN 978-7-111-62483-7

责任编辑 / 卢璐

封面设计 / Randy Lomer, 张健

出版发行 / 机械工业出版社

地 址 / 北京市西城区百万庄大街 22 号 (邮政编码 100037)

印 刷 / 北京瑞德印刷有限公司

开 本 / 178 毫米 × 233 毫米 16 开本 16.25 印张

版 次 / 2019 年 5 月第 1 版 2019 年 6 月第 2 次印刷

定 价 / 89.00 元 (册)

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88379426; 88361066

购书热线: (010) 68326294

投稿热线: (010) 88379604

读者信箱: hzit@hzbook.com

# O'Reilly Media, Inc. 介绍

O'Reilly Media 通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自 1978 年开始，O'Reilly 一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly 的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly 为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了 Make 杂志，从而成为 DIY 革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly 的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly 现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版，在线服务或者面授课程，每一项 O'Reilly 的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

## 业界评论

“O'Reilly Radar 博客有口皆碑。”

——Wired

“O'Reilly 凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference 是聚集关键思想领袖的绝对典范。”

——CRN

“一本 O'Reilly 的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim 是位特立独行的商人，他不光放眼于最长远、最广阔的视野并且切实地按照 Yogi Berra 的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去 Tim 似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

# 译者序

深度神经网络是一种层数更多、规模更大的人工神经网络，较传统神经网络在处理能力上有大幅的提升。2006年，加拿大多伦多大学的教授杰弗里·辛顿（Geoffrey Hinton）在深度信念网络方面进行了卓越的工作，开辟了深度学习这个新的技术领域。目前，深度学习技术已经成为新一代人工智能技术的研究与开发热点，得到了全球的普遍关注，每天都有相关的报道，每年有大量的论文发表，不断刷新着语音识别、图像分类、商品推荐等各应用领域智能处理水平的纪录。与此同时，深度学习模型难以解释、参数调优困难、参数规模大、训练周期长等问题也困扰着研究和开发人员。

如何让深度学习模型设计更加简洁高效，如何处理模型参数调试中遇到的困扰和难题，如何将深度学习快速地应用到具体的业务领域，这些都是深度学习技术研究与开发者需要掌握的内容。本书作为一本聚焦深度学习实际应用的开发指南，很好地解决了这些问题。本书的作者是一位资深的软件工程师，有着丰富的软件开发和调试经验。本书记录了作者从实际工作中总结出来的很多开发技巧，非常适合开发实际应用的深度学习工程师阅读和参考。

本书的第1章从深度学习相关的基本概念开始，介绍了典型的神经网络结构和各种层的设计特点，然后对深度学习中常见的数据集进行了介绍，最后对数据预处理和数据集的划分进行了细致的阐述。第2章是与深度神经网络调试相关的通用技巧，主要涉及如何解决遇到的问题，包括排查错误、检查结果、选择激活函数、正则化和Dropout、设置训练参数等技巧。第3~15章以实际例子，介绍了深度学习在文本处理、图像处理、音乐处理等方面的技巧，涵盖了深度学习主要应用的领域和数据类型，内容非常丰富。最后一章作者从实际使用的角度告诉读者如何在生产系统中部署机器学习应用，使得本书的内容更加贴近实际。

深度学习的技术还处在不断迭代更新的阶段，每天都有新的研究进展发布，新的开发工具开源，以及新的技术挑战出现。本书的内容是当前深度学习设计开发技巧的总结，需要读者在实践中不断尝试，进而提升自己的技术水平，这样才会不断加深对深度学习技术的理解和把握，创造出更加优秀的算法、模型和应用。希望读者朋友在深度学习的实践中不断总结提炼，贡献出更多优秀的图书作品。

每一次翻译工作都是一次难忘的学习之旅，我们非常珍惜这个机会。非常感谢本书的作者和机械工业出版社华章公司的编辑，是他们辛勤的工作为我们创造了难得的机会，让我们能够和广大读者一起走进深度学习的世界，领略新一代人工智能技术的风采。这里还要感谢公司同事和家人的大力支持，他们的鼓励给了我们不断前进的动力。本书翻译过程中，我们努力表达作者的真知灼见，但因水平有限，难免有词不达意的地方和疏漏之处，敬请读者朋友不吝赐教。

译者

2018年12月

# 译者简介

**李君婷** 国家电网有限公司信息通信分公司，主要从事电力信息通信运维数据统计分析、项目管理等工作，研究兴趣包括深度学习、数据科学、颠覆性创新等。

**闫龙川** 国家电网有限公司信息通信分公司，主要从事电力信息通信技术研究工作，研究兴趣包括深度学习、强化学习、云计算、数据中心等。

**俞学豪** 国家电网有限公司信息通信分公司，主要从事电力信息通信技术研究与管理工  
作，研究兴趣包括人工智能、云计算、绿色数据中心等。

**高德荃** 国家电网有限公司信息通信分公司，主要从事电力信息通信技术研究工作，研究兴趣包括数据科学与人工智能、地理空间分析等。

## 作者介绍

---

Douwe Osinga 曾供职于 Google，是一位经验丰富的工程师、环球旅行者和三个初创企业的创始人。他的流行软件项目网站 (<https://douweosinga.com/projects>) 包括了机器学习在内的多个有趣的领域。

## 封面介绍

---

本书封面上的动物是一只潜鸟，或称北美食鱼大鸟（学名：普通潜鸟），在美国北部和加拿大的偏远淡水湖泊附近，以及格陵兰岛的南部地区、冰岛、挪威和阿拉斯加可以看到这种鸟类。

在夏季繁殖季节，成年潜鸟的羽毛呈现出高贵的气质。它的头部和颈部是黑色的，带有彩虹般的光泽，背部是黑色和白色斑点，胸部为白色。在冬季和迁徙过程中，潜鸟的背部羽毛变成浅灰色，喉咙羽毛是白色的。潜鸟是季节性的一夫一妻制，在繁殖季节成对在一起，在冬季迁徙时分开。雌性潜鸟每年产两粒卵。幼鸟在孵化后 1~2 天离开巢穴，在 10~11 周能够飞行。

潜鸟蹼状脚长在身体靠后的地方，因此是强大的游泳者，但这种遗传性的变异阻碍了它在陆地上的活动。在几乎没有溅起水花的情况下，它就可以滑到水面下觅食。它的食物主要是小鱼，偶尔有甲壳动物或青蛙。它在觅食时单独行动，在夜间则聚集成群。

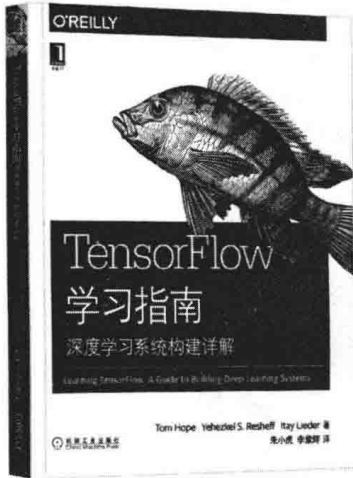
潜鸟是北美荒野的象征，它的叫声是北方森林初夏的一种典型声音。

O'Reilly 封面上的许多动物都已濒临灭绝，但它们对世界来说都很重要。如果你想要了解如何帮助他们，请访问 [animals.oreilly.com](http://animals.oreilly.com)。

本书封面图片来自《英国鸟类》。



# 推荐阅读

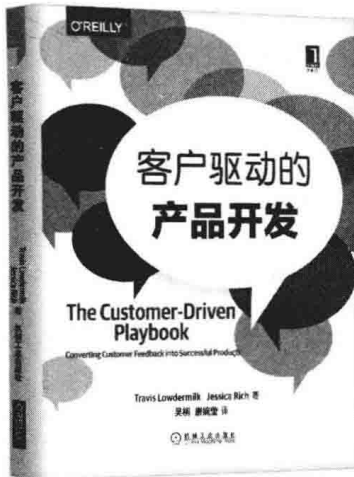


## TensorFlow学习指南: 深度学习系统构建详解

作者: Tom Hope ISBN: 978-7-111-60072-5 定价: 69.00元

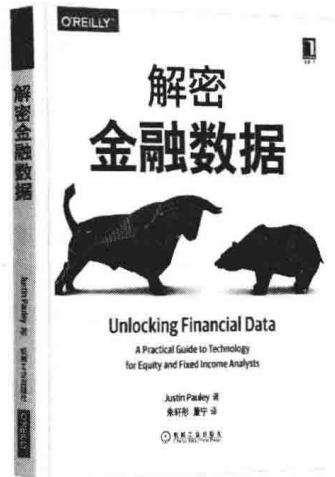
## 机器学习实战: 基于Scikit-Learn和TensorFlow

作者: Aurélien Géron 等 ISBN: 978-7-111-60302-3 定价: 119.00元



## 客户驱动的产品开发

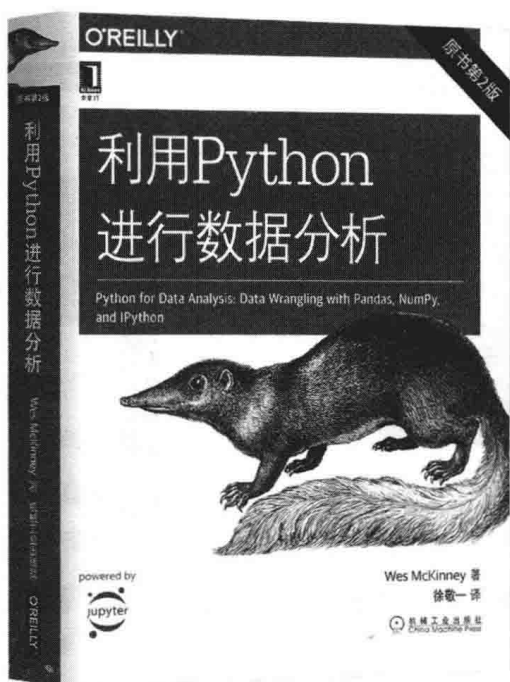
作者: Travis Lowdermilk, Jessica Rich ISBN: 978-7-111-60442-6 定价: 69.00元



## 解密金融数据

作者: Justin Pauley ISBN: 978-7-111-60788-5 定价: 79.00元

## 推荐阅读



### 利用Python进行数据分析（原书第2版）

书号：978-7-111-60370-2 作者：Wes McKinney 定价：119.00元

Python数据分析经典畅销书全新升级，第1版中文版累计印刷10万册  
Python pandas创始人亲自执笔，Python语言的核心开发人员鼎力推荐  
针对Python 3.6进行全面修订和更新，涵盖新版的pandas、NumPy、IPython和Jupyter，并  
增加大量实际案例，可以帮助你高效解决一系列数据分析问题

# 目录

前言 .....	1
<b>第 1 章 工具与技术 .....</b>	<b>9</b>
1.1 神经网络的类型 .....	9
1.2 数据获取 .....	19
1.3 数据预处理 .....	27
<b>第 2 章 摆脱困境 .....</b>	<b>34</b>
2.1 确定我们遇到的问题 .....	34
2.2 解决运行过程中的错误 .....	36
2.3 检查中间结果 .....	38
2.4 为最后一层选择正确的激活函数 .....	39
2.5 正则化和 Dropout .....	40
2.6 网络结构、批尺寸和学习率 .....	42
<b>第 3 章 使用词嵌入计算文本相似性 .....</b>	<b>44</b>
3.1 使用预训练的词嵌入发现词的相似性 .....	45
3.2 Word2vec 数学特性 .....	47
3.3 可视化词嵌入 .....	49
3.4 在词嵌入中发现实体类 .....	51
3.5 计算类内部的语义距离 .....	55
3.6 在地图上可视化国家数据 .....	57
<b>第 4 章 基于维基百科外部链接构建推荐系统 .....</b>	<b>58</b>
4.1 收集数据 .....	58
4.2 训练电影嵌入 .....	62

4.3 构建电影推荐系统.....	66
4.4 预测简单的电影属性.....	67
<b>第 5 章 按照示例文本的风格生成文本.....</b>	<b>69</b>
5.1 获取公开领域书籍文本.....	69
5.2 生成类似莎士比亚的文本.....	70
5.3 使用 RNN 编写代码.....	74
5.4 控制输出温度.....	76
5.5 可视化循环神经网络的活跃程度.....	78
<b>第 6 章 问题匹配.....</b>	<b>80</b>
6.1 从 Stack Exchange 网站获取数据.....	80
6.2 使用 Pandas 探索数据.....	82
6.3 使用 Keras 对文本进行特征化.....	83
6.4 构建问答模型.....	84
6.5 用 Pandas 训练模型.....	86
6.6 检查相似性.....	88
<b>第 7 章 推荐表情符号.....</b>	<b>90</b>
7.1 构建一个简单的情感分类器.....	90
7.2 检验一个简单的分类器.....	93
7.3 使用卷积网络进行情感分析.....	95
7.4 收集 Twitter 数据.....	97
7.5 一个简单的表情符号预测器.....	99
7.6 Dropout 和多层窗口.....	100
7.7 构建单词级模型.....	102
7.8 构建你自己的嵌入.....	104
7.9 使用循环神经网络进行分类.....	106
7.10 可视化一致性 / 不一致性.....	108
7.11 组合模型.....	111
<b>第 8 章 Sequence-to-Sequence 映射.....</b>	<b>113</b>
8.1 训练一个简单的 Sequence-to-Sequence 模型.....	113
8.2 从文本中提取对话.....	115
8.3 处理开放词汇表.....	117
8.4 训练 seq2seq 聊天机器人.....	119

<b>第 9 章 复用预训练的图像识别网络 .....</b>	<b>123</b>
9.1 加载预训练网络 .....	124
9.2 图像预处理 .....	124
9.3 推测图像内容 .....	126
9.4 使用 Flickr API 收集一组带标签的图像 .....	128
9.5 构建一个分辨猫狗的分类器 .....	129
9.6 改进搜索结果 .....	131
9.7 复训图像识别网络 .....	133
<b>第 10 章 构建反向图像搜索服务 .....</b>	<b>137</b>
10.1 从维基百科中获取图像 .....	137
10.2 向 $N$ 维空间投影图像 .....	140
10.3 在高维空间中寻找最近邻 .....	141
10.4 探索嵌入中的局部邻域 .....	143
<b>第 11 章 检测多幅图像 .....</b>	<b>145</b>
11.1 使用预训练的分类器检测多个图像 .....	145
11.2 使用 Faster RCNN 进行目标检测 .....	149
11.3 在自己的图像上运行 Faster RCNN .....	152
<b>第 12 章 图像风格 .....</b>	<b>155</b>
12.1 可视化卷积神经网络激活值 .....	156
12.2 尺度和缩放 .....	159
12.3 可视化神经网络所见 .....	161
12.4 捕捉图像风格 .....	164
12.5 改进损失函数以提升图像相干性 .....	168
12.6 将风格迁移至不同图像 .....	169
12.7 风格内插 .....	171
<b>第 13 章 用自编码器生成图像 .....</b>	<b>173</b>
13.1 从 Google Quick Draw 中导入绘图 .....	174
13.2 为图像创建自编码器 .....	176
13.3 可视化自编码器结果 .....	178
13.4 从正确的分布中采样图像 .....	180
13.5 可视化变分自编码器空间 .....	183
13.6 条件变分编码器 .....	185

<b>第 14 章 使用深度网络生成图标 .....</b>	<b>189</b>
14.1 获得训练用的图标 .....	190
14.2 将图标转换为张量表示 .....	193
14.3 使用变分自编码器生成图标 .....	194
14.4 使用数据扩充提升自编码器的性能 .....	196
14.5 构建生成式对抗网络 .....	198
14.6 训练生成式对抗网络 .....	200
14.7 显示 GAN 生成的图标 .....	202
14.8 将图标编码成绘图指令 .....	204
14.9 训练 RNN 绘制图标 .....	205
14.10 使用 RNN 生成图标 .....	207
<b>第 15 章 音乐与深度学习 .....</b>	<b>210</b>
15.1 为音乐分类器创建训练数据集 .....	211
15.2 训练音乐风格检测器 .....	213
15.3 对混淆情况进行可视化 .....	215
15.4 为已有的音乐编制索引 .....	217
15.5 设置 Spotify API .....	219
15.6 从 Spotify 中收集播放列表和歌曲 .....	221
15.7 训练音乐推荐系统 .....	224
15.8 使用 Word2vec 模型推荐歌曲 .....	225
<b>第 16 章 生产化部署机器学习系统 .....</b>	<b>228</b>
16.1 使用 scikit-learn 最近邻计算嵌入 .....	229
16.2 使用 Postgres 存储嵌入 .....	230
16.3 填充和查询 Postgres 存储的嵌入 .....	231
16.4 在 Postgres 中存储高维模型 .....	233
16.5 使用 Python 编写微服务 .....	234
16.6 使用微服务部署 Keras 模型 .....	236
16.7 从 Web 框架中调用微服务 .....	237
16.8 Tensorflow seq2seq 模型 .....	238
16.9 在浏览器中执行深度学习模型 .....	240
16.10 使用 TensorFlow 服务执行 Keras 模型 .....	243
16.11 在 iOS 中使用 Keras 模型 .....	245

## 深度学习简史

当前深度学习的热潮，其根源可令人惊讶地追溯到 20 世纪 50 年代。虽然“智能机器”的模糊概念可以进一步追溯到更早期的科幻小说和各类科学设想中，但是到了 20 世纪 50 年代和 60 年代，才真正出现了“人工神经网络”的最初理念，该理念基于有关生物神经元的极简模型。在这些模型中，由 Frank Rosenblatt 提出的感知机系统引起了大家的极大兴趣。通过连接到一个简单的“照相机”回路，它可以学会区分不同类型的物体。该系统的第一个版本以软件形式在 IBM 计算机上运行，但是其后续的版本都是用纯硬件来实现的。

对多层感知机 (Multilayer Perceptron, MLP) 模型的兴趣在 20 世纪 60 年代持续不断。到了 1969 年，Marvin Minsky 和 Seymour Papert 出版了《感知机》(Perceptrons, MIT 出版社) 一书，彻底改变了这种形势。这本书证明了线性感知器不能对非线性函数 (XOR) 的行为进行分类。虽然该证明存在局限性 (该书出版时，非线性感知器模型已被提出，作者也注意到了这一问题)，但是该书的出版预示着神经网络模型研究基金的急剧减少。直到 20 世纪 80 年代，随着新一代研究人员的逐步崛起，相关研究才得以恢复。

随着计算能力的提高以及反向传播 (back-propagation) 技术的发展 (反向传播技术自 20 世纪 60 年代被提出以来，有很多种不同的形式，但是直到 80 年代才开始普遍应用)，人们对神经网络重新产生了兴趣。不仅计算机拥有了训练更大型网络的能力，我们也拥有了有效训练更深层网络的能力。最初的卷

积神经网络将这些技术发展与哺乳动物大脑的视觉识别模型结合起来，首次产生了能够有效地识别诸如手写数字和人脸等复杂图像的网络。卷积网络通过将相同的“子网络”应用到图像的不同位置并将这些结果聚合到更高级的特征中来实现这一点。在本书第 12 章中，我们会详细介绍这方面的内容。

20 世纪 90 年代和 21 世纪 00 年代初期，随着支持向量机 (SVM) 和决策树等更“易于理解”的模型变得流行，人们对神经网络的兴趣再次出现了下降。对于当时许多数据源来说，SVM 被证明是非常优秀的分类器，特别是在与人工特征相结合时更是如此。在计算机视觉中，“特征工程”开始变得流行起来。该技术涉及为图片中的小元素构建特征检测器，并人工将其组合成能够识别更复杂形态的模型。后来科研人员发现，深度学习网络也能够学会识别类似的特征，并能够学会以非常相似的方式组合这些特征。在第 12 章中，我们将会探讨这些网络的内部工作机制，并对深度学习网络所学的内容进行可视化。

21 世纪 00 年代后期，随着图形处理单元 (GPU) 通用编程的出现，神经网络架构在竞争中取得了长足的进步。GPU 包含数千个微处理器，它们可以每秒并行执行数万亿次操作。GPU 最初是为计算机游戏开发的，主要目的是实现复杂 3D 场景实时渲染，事实证明，GPU 也能够用于并行训练神经网络，可以实现 10 倍或更高倍数的计算速度提升。

另外一件促使深度学习领域长足进步的事情，是互联网的发展为其提供了大量可用的训练数据。以往研究人员只能使用数千幅图像训练分类器，现在已经可以提供几千万甚至数以亿计的图像了。结合更大型的网络，神经网络技术现在迎来了绽放光芒的机会。这种优势仅在最近几年才开始持续出现，伴随着技术改进与现实应用，神经网络技术逐步应用到了图像识别之外的很多领域，包括机器翻译、语音识别和图像合成。

## 为什么是现在

计算能力和改进技术的爆发式发展，使人们对于神经网络的兴趣逐步增加，同时我们也看到了神经网络技术在可用性上取得了巨大的进步。特别是像 TensorFlow、Theano 和 Torch 这样的深度学习框架使得非专业人士也能够构建出复杂的神经网络来解决自己的机器学习问题。这使得以往需要数月甚至



数年手动编码和辛勤付出（编写高效的 GPU 内核真的十分困难！）的任务，转变成任何一个人都可以在一下午（或者几天内）完成的任务。可用性的提升极大地增加了有能力开展深度学习问题研究的人员数量。正如本书在后续章节中将展示的那样，具有更高抽象级别的框架，比如 Keras，使得任何具有 Python 和相关工具知识的人都能够开展一些有趣的实验。

回答“为什么是现在”这个问题的第二个要素是，每个人都可以使用大型的数据集。是的，Facebook 和 Google 在访问数十亿图片、用户评论方面仍占据优势，但是你也可以通过各种数据源获得包含数百万条目的数据集。在第 1 章中，我们将会探讨各种可选的数据源，并且纵览全书，每章的示例代码通常都会在各章开头向你展示如何获得所需的训练数据。

与此同时，私营公司也开始生产和收集更大数量级的数据，这使得整个深度学习领域在商业上忽然变得越来越令人着迷。一个能辨别猫和狗的模型已经非常不错了，然而一个能够在使用所有历史销售数据的基础上将销售额提升 15% 的模型，对于一个公司来说则可能意味着生死存亡。

## 你需要知道什么

如今，对于深度学习来说，有很多平台、技术和编程语言可供选择。在本书中，所有的例子都是用 Python 编写的，并且大部分代码的实现都依赖于优秀的 Keras 框架。本书的示例代码可以在 GitHub 上的 Python notebook 中找到，每章的代码存放在一个 notebook 中。因此，对读者来说，具备以下知识技能将有助于阅读本书：

### *Python*

Python 3 是首选版本，你也可以使用 Python 2.7。我们会用到各类 helper 库，你可使用 pip 轻松安装它们。本书涉及的代码大都比较简单易懂，所以即使是一名新手也可以跟着本书进行实践。

### *Keras*

机器学习的繁重工作几乎全部是由 Keras 完成的。Keras 是对 TensorFlow 或 Theano 深度学习框架的抽象封装。Keras 能够轻松地使用可读的方式