

O'REILLY®  
オライリー・ジャパン

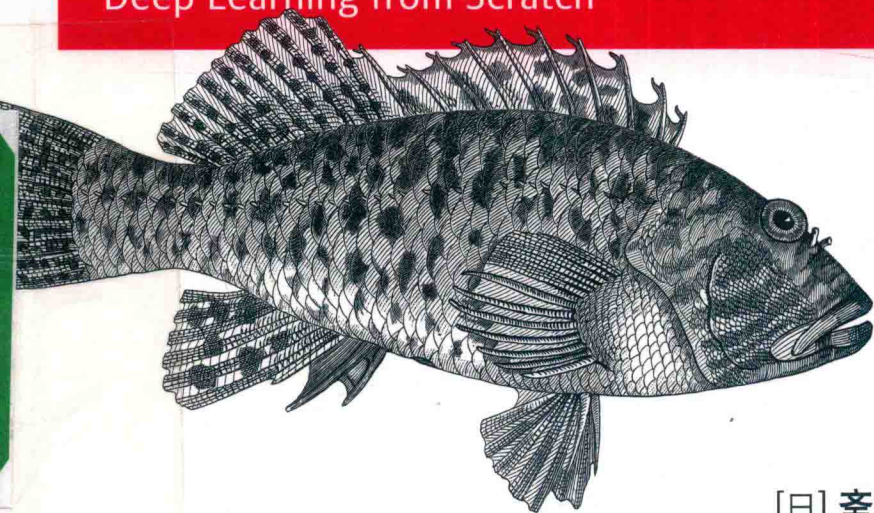
TURING

图灵程序设计丛书

# 深度学习入门

基于Python的理论与实现

Deep Learning from Scratch



[日] 斋藤康毅 著  
陆宇杰 译



中国工信出版集团



人民邮电出版社  
POSTS & TELECOM PRESS

TURING

图灵程序设计丛书



# 深度学习入门

## 基于Python的理论与实现

Deep Learning from Scratch

[日] 斋藤康毅 著  
陆宇杰 译



O'REILLY®

Beijing • Boston • Farnham • Sebastopol • Tokyo

O'Reilly Japan, Inc. 授权人民邮电出版社出版

人民邮电出版社  
北京

## 图书在版编目 (C I P) 数据

深度学习入门：基于Python的理论与实现 / (日) 斋藤康毅著；陆宇杰译. — 北京：人民邮电出版社，2018.7 (2018.8重印)

(图灵程序设计丛书)

ISBN 978-7-115-48558-8

I. ①深… II. ①斋… ②陆… 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112509号

### 内 容 提 要

本书是深度学习真正意义上的入门书，深入浅出地剖析了深度学习的原理和相关技术。书中使用Python 3，尽量不依赖外部库或工具，带领读者从零创建一个经典的深度学习网络，使读者在此过程中逐步理解深度学习。书中不仅介绍了深度学习和神经网络的概念、特征等基础知识，对误差反向传播法、卷积神经网络等也有深入讲解，此外还介绍了学习相关的实用技巧，自动驾驶、图像生成、强化学习等方面的应用，以及为什么加深层可以提高识别精度等疑难问题。

本书适合深度学习初学者阅读，也可作为高校教材使用。

- 
- ◆ 著 [日] 斋藤康毅
  - 译 陆宇杰
  - 责任编辑 杜晓静
  - 执行编辑 刘香娣
  - 责任印制 周昇亮
  - ◆ 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：880×1230 1/32
  - 印张：9.625
  - 字数：300千字 2018年7月第1版
  - 印数：4 001-7 000册 2018年8月北京第2次印刷
  - 著作权合同登记号 图字：01-2017-0526号

---

定价：59.00元

读者服务热线：(010)51095186转600 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

## 作者介绍

### 斋藤康毅

日本东京工业大学毕业，并完成东京大学研究生院课程。现从事计算机视觉与机器学习相关的研究和开发工作。是 *Introducing Python*、*Python in Practice*、*The Elements of Computing Systems*、*Building Machine Learning Systems with Python* 的日文版译者。

## 译者介绍

### 陆宇杰

众安科技NLP算法工程师。主要研究方向为自然语言处理及其应用，对图像识别、机器学习、深度学习等领域有密切关注。Python爱好者。

站在巨人的肩上  
**Standing on Shoulders of Giants**



iTuring.cn

试读结束：需要全本请在线购买：[www.ertongbook.com](http://www.ertongbook.com)

# 版权声明

Copyright © 2016 Koki Saitoh, O'Reilly Japan, Inc.

Posts and Telecommunications Press, 2018.

Authorized translation of the Japanese edition of “Deep Learning from Scratch” © 2016 O'Reilly Japan, Inc. This translation is published and sold by permission of O'Reilly Japan, Inc., the owner of all rights to publish and sell the same.

日文原版由 O'Reilly Japan, Inc. 出版，2016。

简体中文版由人民邮电出版社出版，2018。日文原版的翻译得到 O'Reilly Japan, 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 Japan, 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。

# 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：‘如果你在路口遇到岔路口，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

深度学习的浪潮已经汹涌澎湃了一段时间了，市面上相关的图书也已经出版了很多。其中，既有知名学者伊恩·古德费洛(Ian Goodfellow)等人撰写的系统介绍深度学习基本理论的《深度学习》，也有各种介绍深度学习框架的使用方法的入门书。你可能会问，现在再出一本关于深度学习的书，是不是“为时已晚”？其实并非如此，因为本书考察深度学习的角度非常独特，它的出版可以说是“千呼万唤始出来”。

本书最大的特点是“剖解”了深度学习的底层技术。正如美国物理学家理查德·费曼(Richard Phillips Feynman)所说：“What I cannot create, I do not understand.”只有创造一个东西，才算真正弄懂了一个问题。本书就是教你如何创建深度学习模型的一本书。并且，本书不使用任何现有的深度学习框架，尽可能仅使用最基本的数学知识和Python库，从零讲解深度学习核心问题的数学原理，从零创建一个经典的深度学习网络。

本书的日文版曾一度占据了东京大学校内书店(本乡校区)理工类图书的畅销书榜首。各类读者阅读本书，均可有所受益。对于非AI方向的技术人员，本书将大大降低入门深度学习的门槛；对于在校的大学生、研究生，本书不失为学习深度学习的一本好教材；即便是对于在工作中已经熟练使用框架开发各类深度学习模型的读者，也可以从本书中获得新的体会。

本书从开始翻译到出版，前前后后历时一年之久。译者翻译时力求忠于原文，表达简练。为了保证翻译质量，每翻译完一章后，译者都会放置一段



时间，再重新检查一遍。图灵公司的专业编辑们又进一步对译稿进行了全面细致的校对，提出了许多宝贵意见，在此表示感谢。但是，由于译者才疏学浅，书中难免存在一些错误或疏漏，恳请读者批评指正，以便我们在重印时改正。

最后，希望本书的出版能为国内的AI技术社区添砖加瓦！

陆宇杰

2018年2月 上海

# 前言

科幻电影般的世界已经变成了现实——人工智能战胜过日本将棋、国际象棋的冠军，最近甚至又打败了围棋冠军；智能手机不仅可以理解人们说的话，还能在视频通话中进行实时的“机器翻译”；配备了摄像头的“自动防撞的车”保护着人们的生命安全，自动驾驶技术的实用化也为期不远。环顾我们的四周，原来被认为只有人类才能做到的事情，现在人工智能都能毫无差错地完成，甚至试图超越人类。因为人工智能的发展，我们所处的世界正在逐渐变成一个崭新的世界。

在这个发展速度惊人的世界背后，深度学习技术在发挥着重要作用。对于深度学习，世界各地的研究人员不吝褒奖之辞，称赞其为革新性技术，甚至有人认为它是几十年才有一次的突破。实际上，深度学习这个词经常出现在报纸和杂志中，备受关注，就连一般大众也都有所耳闻。

本书就是一本以深度学习为主题的书，目的是让读者尽可能深入地理解深度学习的技术。因此，本书提出了“从零开始”这个概念。

本书的特点是通过实现深度学习的过程，来逼近深度学习的本质。通过实现深度学习的程序，尽可能无遗漏地介绍深度学习相关的技术。另外，本书还提供了实际可运行的程序，供读者自己进行各种各样的实验。

为了实现深度学习，我们需要经历很多考验，花费很长时间，但是相应地也能学到和发现很多东西。而且，实现深度学习的过程是一个有趣的、令

人兴奋的过程。希望读者通过这一过程可以熟悉深度学习中使用技术，并能从中感受到快乐。

目前，深度学习活跃在世界上各个地方。在几乎人手一部的智能手机中、开启自动驾驶的汽车中、为Web服务提供动力的服务器中，深度学习都在发挥着作用。此时此刻，就在很多人没有注意到的地方，深度学习正在默默地发挥着其功能。今后，深度学习势必更加活跃。为了让读者理解深度学习的相关技术，感受到深度学习的魅力，笔者写下了本书。

## 本书的理念

本书是一本讲解深度学习的书，将从最基础的内容开始讲起，逐一介绍理解深度学习所需的知识。书中尽可能用平实的语言来介绍深度学习的概念、特征、工作原理等内容。不过，本书并不是只介绍技术的概要，而是旨在让读者更深入地理解深度学习。这是本书的特色之一。

那么，怎样才能更深入地理解深度学习呢？在笔者看来，最好的办法就是亲自实现。从零开始编写可实际运行的程序，一边看源代码，一边思考。笔者坚信，这种做法对正确理解深度学习（以及那些看上去很高级的技术）是很重要的。这里用了“从零开始”一词，表示我们将尽可能地不依赖外部的现成品（库、工具等）。也就是说，本书的目标是，尽量不使用内容不明的黑盒，而是从自己能理解的最基础的知识出发，一步一步地实现最先进的深度学习技术。并通过这一实现过程，使读者加深对深度学习的理解。

如果把本书比作一本关于汽车的书，那么本书并不会教你怎么开车，其着眼点不是汽车的驾驶方法，而是要让读者理解汽车的原理。为了让读者理解汽车的结构，必须打开汽车的引擎盖，把零件一个一个地拿在手里观察，并尝试操作它们。之后，用尽可能简单的形式提取汽车的本质，并组装汽车模型。本书的目标是，通过制造汽车模型的过程，让读者感受到自己可以实际制造出汽车，并在这一过程中熟悉汽车相关的技术。

为了实现深度学习，本书使用了Python这一编程语言。Python非常受欢迎，初学者也能轻松使用。Python尤其适合用来制作样品（原型），使用

Python可以立刻尝试突然想到的东西，一边观察结果，一边进行各种各样的实验。本书将在讲解深度学习理论的同时，使用Python实现程序，进行各种实验。



在光看数学式和理论说明无法理解的情况下，可以尝试阅读源代码并运行，很多时候思路都会变得清晰起来。对数学式感到困惑时，就阅读源代码来理解技术的流程，这样的事情相信很多人都经历过。本书通过实际实现（落实到代码）来理解深度学习，是一本强调“工程”的书。书中会出现很多数学式，但同时也会有很多程序员视角的源代码。

## 本书面向的读者

本书旨在让读者通过实际动手操作来深入理解深度学习。为了明确本书的读者对象，这里将本书涉及的内容列举如下。

- 使用Python，尽可能少地使用外部库，从零开始实现深度学习的程序。
- 为了让Python的初学者也能理解，介绍Python的使用方法。
- 提供实际可运行的Python源代码，同时提供可以让读者亲自实验的学习环境。
- 从简单的机器学习问题开始，最终实现一个能高精度地识别图像的系统。
- 以简明易懂的方式讲解深度学习和神经网络的理论。
- 对于误差反向传播法、卷积运算等乍一看很复杂的技术，使读者能够在实现层面上理解。
- 介绍一些学习深度学习时有用的实践技巧，如确定学习率的方法、权重的初始值等。
- 介绍最近流行的Batch Normalization、Dropout、Adam等，并进行实现。
- 讨论为什么深度学习表现优异、为什么加深层能提高识别精度、为什么隐藏层很重要等问题。
- 介绍自动驾驶、图像生成、强化学习等深度学习的应用案例。

## 本书不面向的读者

明确本书不适合什么样的读者也很重要。为此，这里将本书不会涉及的内容列举如下。

- 不介绍深度学习相关的最新研究进展。
- 不介绍Caffe、TensorFlow、Chainer等深度学习框架的使用方法。
- 不介绍深度学习的详细理论，特别是神经网络相关的详细理论。
- 不详细介绍用于提高识别精度的参数调优相关的内容。
- 不会为了实现深度学习的高速化而进行GPU相关的实现。
- 本书以图像识别为主题，不涉及自然语言处理或者语音识别的例子。

综上，本书不涉及最新研究和理论细节。但是，读完本书之后，读者应该有能力进一步去阅读最新的论文或者神经网络相关的理论方面的技术书。



本书以图像识别为主题，主要学习使用深度学习进行图像识别时所需的技术。自然语言处理或者语音识别等不是本书的讨论对象。

## 本书的阅读方法

学习新知识时，只听别人讲解的话，有时会无法理解，或者会立刻忘记。正如“不闻不若闻之，闻之不若见之，见之不若知之，知之不若行之”<sup>①</sup>，在学习新东西时，没有什么比实践更重要了。本书在介绍某个主题时，都细心地准备了一个可以实践的场所以——能够作为程序运行的源代码。

本书会提供Python源代码，读者可以自己动手实际运行这些源代码。在阅读源代码的同时，可以尝试去实现一些自己想到的东西，以确保真正

---

① 出自荀子《儒效篇》。

理解了。另外，读者也可以使用本书的源代码，尝试进行各种实验，反复试错。

本书将沿着“理论说明”和“Python实现”两个路线前进。因此，建议读者准备好编程环境。本书可以使用Windows、Mac、Linux中的任何一个系统。关于Python的安装和使用方法将在第1章介绍。另外，本书中用到的程序可以从以下网址下载。

<http://www.ituring.com.cn/book/1921>

## 让我们开始吧

通过前面的介绍，希望读者了解本书大概要讲的内容，产生继续阅读的兴趣。

最近出现了很多深度学习相关的库，任何人都可以方便地使用。实际上，使用这些库的话，可以轻松地运行深度学习的程序。那么，为什么我们还要特意花时间从零开始实现深度学习呢？一个理由就是，在制作东西的过程中可以学到很多。

在制作东西的过程中，会进行各种各样的实验，有时也会卡住，抱着脑袋想为什么会这样。这种费时的工作对深刻理解技术而言是宝贵的财富。像这样认真花费时间获得的知识在使用现有的库、阅读最新的文章、创建原创的系统时都大有用处。而且最重要的是，制作本身就是一件快乐的事情。（还需要快乐以外的其他什么理由吗？）

既然一切都准备好了，下面就让我们踏上实现深度学习的旅途吧！

## 表述规则

本书在表述上采用如下规则。

### 粗体字 (Bold)

用来表示新引入的术语、强调的要点以及关键短语。

### 等宽字 (Constant Width)

用来表示下面这些信息：程序代码、命令、序列、组成元素、语句选项、分支、变量、属性、键值、函数、类型、类、命名空间、方法、模块、属性、参数、值、对象、事件、事件处理器、XML 标签、HTML 标签、宏、文件的内容、来自命令行的输出等。若在其他地方引用了以上这些内容(如变量、函数、关键字等)，也会使用该格式标记。

### 等宽粗体字 (Constant Width Bold)

用来表示用户输入的命令或文本信息。在强调代码的作用时也会使用该格式标记。

### 等宽斜体字 (Constant Width Italic)

用来表示必须根据用户环境替换的字符串。



用来表示提示、启发以及某些值得深究的内容的补充信息。



表示程序库中存在的 bug 或时常会发生的问题等警告信息，引起读者对该处内容的注意。

## 读者意见与咨询

虽然笔者已经尽最大努力对本书的内容进行了验证与确认，但仍不免在某些地方出现错误或者容易引起误解的表达等，给读者的理解带来困扰。如果读者遇到这些问题，请及时告知，我们在本书重印时会将其改正，在此先表示不胜感激。与此同时，也希望读者能够为本书将来的修订提出中肯的建议。本书编辑部的联系方式如下。



株式会社 O'Reilly Japan

电子邮件 [japan@oreilly.co.jp](mailto:japan@oreilly.co.jp)

本书的主页地址如下。

<http://www.ituring.com.cn/book/1583>

<http://www.oreilly.co.jp/books/9784873117584> (日语)

<https://github.com/oreilly-japan/deep-learning-from-scratch>

关于 O'Reilly 的其他信息，可以访问下面的 O'Reilly 主页查看。

<http://www.oreilly.com/> (英语)

<http://www.oreilly.co.jp/> (日语)

## 致谢

首先，笔者要感谢推动了深度学习相关技术(机器学习、计算机科学等)发展的研究人员和工程师。本书的完成离不开他们的研究工作。其次，笔者还要感谢在图书或网站上公开有用信息的各位同仁。其中，斯坦福大学的 CS231n<sup>[5]</sup> 公开课慷慨提供了很多有用的技术和信息，笔者从中学到了很多东西。

在本书执笔过程中，曾受到下列人士的帮助：teamLab 公司的加藤哲朗、喜多慎弥、飞永由夏、中野皓太、中村将达、林辉大、山本辽；Top Studio 公司的武藤健志、增子萌；Flickfit 公司的野村宪司；得克萨斯大学奥斯汀分校 JSPS 海外特别研究员丹野秀崇。他们阅读了本书原稿，提出了很多宝贵的建议，在此深表谢意。另外，需要说明的是，本书中存在的不足或错误均是笔者的责任。

最后，还要感谢 O'Reilly Japan 的宫川直树，在从本书的构想到完成的大约一年半的时间里，宫川先生一直支持着笔者。非常感谢！

2016年9月1日

斋藤康毅

Copyright ©2016 Koki Saitoh, O'Reilly Japan, Inc. All rights reserved.

本书中使用的系统名、产品名都是各公司的商标或注册商标。

正文中有时会省略TM、®、©等标识。

---

株式会社O'Reilly Japan尽最大努力确保了本书内容的正确性，但对运用本书内容所造成的任何结果概不负责，敬请知悉。

试读结束：需要全本请在线购买：[www.ertongbook.com](http://www.ertongbook.com)