

O'REILLY®



图灵程序设计丛书



深度学习 基础与实践

Deep Learning: A Practitioner's Approach



[美] 乔希·帕特森 亚当·吉布森 著
郑明智 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

深度学习基础与实践

Deep Learning: A Practitioner's Approach

[美] 乔希·帕特森 [美] 亚当·吉布森 著
郑明智 译



Beijing • Boston • Farnham • Sebastopol • Tokyo

O'REILLY®

O'Reilly Media, 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 (CIP) 数据

深度学习基础与实践 / (美) 乔希·帕特森
(Josh Patterson), (美) 亚当·吉布森
(Adam Gibson) 著; 郑明智 译. — 北京: 人民邮电
出版社, 2019.7

(图灵程序设计丛书)
ISBN 978-7-115-51542-1

I. ①深… II. ①乔… ②亚… ③郑… III. ①机器学习—研究 IV. ①TP181

中国版本图书馆CIP数据核字(2019)第124039号

内 容 提 要

本书是由两位技术出身的企业管理者编写的深度学习普及书。本书的前四章提供了足够的关于深度学习的理论知识, 包括机器学习的基本概念、神经网络基础、从神经网络到深度网络的演化历程, 以及主流的深度网络架构, 为读者阅读本书剩余内容打下基础。后五章带领读者进行一系列深度学习的实践, 包括建立深层网络、高级调优技术、各种数据类型的向量化和在 Spark 上运行深度学习工作流。

本书适合对深度学习的理论和应用感兴趣的开发人员和研究人员阅读。

-
- ◆ 著 [美] 乔希·帕特森 [美] 亚当·吉布森
译 郑明智
责任编辑 朱巍
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>
北京鑫正大印刷有限公司印刷
 - ◆ 开本: 800×1000 1/16
印张: 24.25
字数: 573千字 2019年7月第1版
印数: 1-3 000册 2019年7月北京第1次印刷
著作权合同登记号 图字: 01-2019-3079号
-

定价: 119.00元

读者服务热线: (010)51095183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

版权声明

© 2017 by Josh Patterson and Adam Gibson.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and Posts & Telecom Press, 2019. Authorized translation of the English edition, 2019 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，2017。

简体中文版由人民邮电出版社出版，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。

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。无论是通过图书出版、在线服务还是面授课程，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 凭借一系列非凡想法（真希望当初我也想到了）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

献给我的儿子Ethan、Griffin和Dane，愿你们勇往直前、坚持不懈。

——乔希·帕特森

前言

本书内容

本书前四章着眼于提供足够的理论和知识，为你（实践者）阅读本书剩余内容打下基础。后五章带领你使用 DL4J 进行一系列深度学习的实践：

- 建立深层网络；
- 高级调优技术；
- 各种数据类型的向量化；
- 在 Spark 上运行深度学习 workflow。



DL4J 是 Deeplearning4j 的缩写

在本书中，DL4J 和 Deeplearning4j 含义相同，指的都是 Deeplearning4j 库的工具套件。

之所以这样组织本书的内容，是因为我们相信存在这样的需求，即一本既涵盖“足够的理论”，同时又提供足够的实践，使得读者能够构建生产级别的深度学习 workflow 的书。我们认为这本书双管齐下，很好地满足了这个需求。

第 1 章回顾了机器学习，尤其是深度学习的概念，以期快速地帮助读者为阅读后续章节打下基础。我们设置这一章，是为了让大量的初学者复习或者了解这些概念，希望本书能够帮助到最广泛的群体。

在此基础上，第 2 章介绍了神经网络的基础知识。这一章的主要内容是神经网络的理论，我们会以易于理解的方式讲述这些知识。第 3 章更进一步，带你快速了解深层网络是如何从神经网络演变而来的。第 4 章介绍了 4 个主流的深层网络架构，并且为阅读本书后续内容提供了基础。

第 5 章通过一系列 Java 代码的例子，带领你运用在本书前半部分学到的技术。第 6 章和

第 7 章阐述了普通神经网络调优的基础知识，以及如何对特定的深层网络架构调优。这两章都与平台无关，适合使用任何深度学习库的实践者阅读。第 8 章回顾了向量化技术，还简要介绍了如何使用 DataVec（DL4J 的 ETL 及向量化 workflow 工具）。第 9 章介绍了如何在 Spark 和 Hadoop 上直接使用 DL4J，展示了 3 个实际的例子，这些例子可以在你的 Spark 集群上运行。

书后提供了多个附录，涵盖了与本书主体内容相关但不甚直接的主题。这些主题包括：

- 人工智能；
- 使用 Maven 管理 DL4J 工程；
- 使用 GPU 进行训练；
- 使用 ND4J API；
- 其他主题。

谁是“实践者”

如今“数据科学”这个词还没有明确的定义，并且经常被用于各种场景。和当今计算机科学中的许多术语一样，数据科学和人工智能（AI）的范围也非常宽广和模糊，这主要是由于机器学习已经扩展到了几乎所有领域。

这种扩展与（20 世纪 90 年代）万维网把 HTML 带到各个行业并把许多新人带入技术领域有相似之处。同样，每天都有各种职业的人——工程师、统计学家、分析师、艺术家——进入机器学习领域。我们编写本书的目标就是为了普及深度学习以及机器学习，让更多人了解它们。

如果你觉得这个主题很有趣，而且正在阅读这个前言，那么你就是实践者，本书就是为你准备的。

目标读者

与那些围绕着小程序的构建来推进的图书不同，本书选择从介绍一系列基础知识开始，带领你完整地走过深度学习的旅程。

我们发现太多的书忽略了那些企业中的实践者经常需要快速复习的核心主题。基于在机器学习领域的经验，我们决定从入门级别的实践者经常需要温习的内容开始，以便更好地支持他们的深度学习工程。

你可能想跳过第 1 章和第 2 章，直奔深度学习基础知识相关章节。但我们希望你不要跳过，因为它们提供了基本的理论和原理，能帮助你顺利地开始学习更加艰深的深度学习主题。我们为不同背景的读者提供了一些阅读建议。

企业中的机器学习实践者

我们把这一类读者细分为以下两组：

- 在职数据科学家；
- Java 工程师。

在职数据科学家

这类读者通常已经建立过模型，而且非常熟悉数据科学领域。如果你属于这一类读者，那么你很可能会跳过第 1 章，并准备快速地浏览第 2 章。我们建议你从第 3 章开始阅读，因为你很可能已经为学习深层网络的基础知识学习做好了准备。

Java 工程师

Java 工程师通常负责将机器学习的代码与生产系统集成。如果你属于这一类读者，我们建议你从第 1 章开始阅读，因为这会帮助你更好地理解数据科学的术语。你应该也会对附录 E 的内容非常感兴趣，因为模型评分代码的集成通常与 ND4J API 直接相关。

企业经营者

我们的一些试读者是财富 500 强企业的经营者，他们对本书内容非常满意，觉得它能帮助自己更好地了解深度学习领域正在发生什么。一位经营者评论说，他大学毕业已有一段时间，而第 1 章对机器学习和深度学习的概念做了很好的回顾。如果你是一位经营者，我们建议你从快速浏览第 1 章开始，这会帮助你回顾一些术语。不过你可能想跳过那些重点讲解 API 和示例代码的章节。

学者

如果你是一名学者，可能想跳过第 1 章和第 2 章，因为在学校已经学过这些内容了。你应该会对神经网络调优以及特定架构调优的相关章节很感兴趣，因为这些内容基于研究，超越了任何具体的深度学习实现。如果你想在 Java 虚拟机 (JVM) 上进行高效的线性代数计算，那么 ND4J 的相关内容你应该也会感兴趣。

排版约定

本书使用了下列排版约定。

- **黑体字**
表示名词和重点强调的内容。
- **等宽字体 (constant width)**
表示程序片段，正文中出现的变量、函数名、数据库、数据类型、环境变量、语句和关键字等程序元素，还用于表示模块名和包名、应该由用户直接输入的命令或其他文本，以及命令的输出。
- **等宽斜体 (constant width italic)**
表示应该被替换为用户提供的值或者由上下文确定的值的文本。



该图标表示提示或建议。



该图标表示普通的注记。



该图标表示警告或警示。

使用示例代码

辅助材料（虚拟机、数据、脚本和自定义命令行工具等）可以从 <https://github.com/deeplearning4j/oreilly-book-dl4j-examples> 下载。

本书的目的在于帮助你完成工作。一般来说，你可以在你的程序和文档中使用本书的代码。只要不是大规模地复制代码，你就不需要联系我们获得授权。举例而言，你写的一个程序用到了本书的几个代码片段，这不需要获得授权。但是，如果你把书中的示例代码刻录到 CD-ROM，并拿去出售和分发，则需要获得授权。在回答问题时引用本书以及本书的示例代码无须获得授权。但如果要在你的产品文档里收录本书中的大量示例代码，则需要获得授权。

欢迎你在使用本书的示例代码时注明出处，但这不是强制要求。通常要注明书名、作者、出版社和 ISBN。例如：*Deep Learning: A Practitioner's Approach* by Josh Patterson and Adam Gibson (O'Reilly). Copyright 2017 Josh Patterson and Adam Gibson, 978-1-4919-1425-0。

如果你认为你对示例代码的使用不在合理使用或上述无须授权的范围之内，那么请通过 permissions@oreilly.com 联系我们。

补充说明

在 Java 代码示例中，我们经常省略导入语句。你可以在实际的代码库中看到完整的导入列表。DL4J、ND4J、DataVec 和更多的 API 信息可在以下网站获得：

<http://deeplearning4j.org/documentation>

你可以在以下网址找到所有代码示例：

<https://github.com/deeplearning4j/oreilly-book-dl4j-examples>

有关 DL4J 系列工具的更多资源，请查看以下网站：

<http://deeplearning4j.org>

O'Reilly Safari



Safari (之前称作 Safari Books Online) 是一个针对企业、政府、教育者和个人的会员制培训和参考平台。

会员可以访问来自 250 多家出版商的上千种图书、培训视频、学习路径、互动式教程和精选播放列表，这些出版商包括 O'Reilly Media、Harvard Business Review、Prentice Hall Professional、Addison-Wesley Professional、Microsoft Press、Sams、Que、Peachpit Press、Adobe、Focal Press、Cisco Press、John Wiley & Sons、Syngress、Morgan Kaufmann、IBM Redbooks、Packt、Adobe Press、FT Press、Apress、Manning、New Riders、McGraw-Hill、Jones & Bartlett、Course Technology 等。

要了解更多信息，可以访问 <http://www.oreilly.com/safari>。

联系我们

请把对本书的评价和问题发给出版社。

美国：

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472

中国：

北京市西城区西直门南大街 2 号成铭大厦 C 座 807 室 (100035)
奥莱利技术咨询 (北京) 有限公司

对于本书的评论和技术性问题，请发送电子邮件到 bookquestions@oreilly.com。如果你发现任何错误、明显的疏漏或困惑之处，或者有任何改进建议，请给 Josh Patterson 发送电子邮件：jpatterson@floe.tv。

O'Reilly 的每一本书都有专属网页，你可以在那儿找到书的相关信息，包括勘误表¹、示例代码以及其他信息。本书的网站地址是：<http://shop.oreilly.com/product/0636920035343.do>

要了解更多 O'Reilly 图书、培训课程、会议和新闻的信息，请访问以下网站：<http://www.oreilly.com>

我们在 Facebook 的地址如下：<http://facebook.com/oreilly>

请关注我们的 Twitter 动态：<http://twitter.com/oreillymedia>

注 1：本书中文版的勘误请到 <http://www.ituring.com.cn/book/2542> 查看和提交。——编者注

我们的 YouTube 视频地址如下：<http://www.youtube.com/oreillymedia>

Josh Patterson 的 Twitter 账号是 @jpatanooga。

致谢

Josh

我邀请了许多比我聪明得多的人来帮助构思和审阅本书内容。像 DL4J 这样规模的项目是无法脱离现实的，在许多社区专家和 Skymind 工程师们的工作的基础上，我才形成了书中的许多想法和准则。

我当时没有想到，与 Adam（在 MLConf 邂逅之后）开发的 DL4J 最终会写成一本书。坦白说，尽管一开始我就参与了 DL4J，但 Adam 做的事情比我多得多。因此，我非常感激 Adam 对这个项目所做的贡献，以及对 JVM 上深度学习想法的贡献，并在前景极不明朗的早期坚持着信念。我想对 Adam 说：是的，你是对的，ND4J 是正确的选择。

写作是一段漫长而孤独的旅程，我要特别感谢 Alex Black 的巨大付出：他不仅审阅了本书，而且还贡献了附录中的内容。Alex 非常了解已出版的神经网络文献，这是编写本书中许多小细节的关键，而且他事无巨细，确保了书中大大小小的知识点都是正确的。没有 Alex Black 的话，第 6 章和第 7 章的内容将至少缩水一半。

Susan Eraly 帮助我们编写了损失函数章节，并且贡献了附录材料。她还订正了书中的许多公式，并且写下了详细的审阅记录。Melanie Warrick 是审阅本书草稿的关键人物，她不但提供了反馈，还为卷积神经网络的内部工作原理编写了注释。

David Kale 是一位勤奋的特设审稿人，他让我在处理许多重要的网络细节和参考文献时小心翼翼。在确定本书的严谨性以及目标读者时，David 提供了学术观点。

每当我就书中内容的取舍发牢骚时，James Long 都是一位严格的听众，并从实践统计学家的角度提出了切实可行的观点。很多时候，对于如何处理复杂的话题并没有明确正确答案，而 James 是与我从多个角度讨论问题的参谋。David Kale 和 Alex Black 经常提醒我注意数学严谨性，而 James 却总是扮演理性反对者的角色，提醒我们要适度，不要“将读者淹没在数学里”。

Vyacheslav “Raver” Kokorin 提升了自然语言处理（NLP）和 Word2Vec 示例代码的质量。

我不会忘记 Skymind 首席执行官 Chris Nicholson 给予我们的支持。Chris 一直支持这本书，正是他给予我们充足的时间和资源，本书才得以完成。

我要感谢那些贡献了附录内容的人：Alex Black（反向传播、DataVec）、Vyacheslav “Raver” Kokorin（GPU）、Susan Eraly（GPU）、Ruben Fiszal（强化学习）。本书在不同阶段的其他审阅者还有 Grant Ingersol、Dean Wampler、Robert Chong、Ted Malaska、Ryan Geno、Lars George、Suneel Marthi、Francois Garillot 和 Don Brown。如果你在书中发现任何错误，责任都在我。

我要感谢备受尊敬的编辑 Tim McGovern 的反馈和批注，以及他对一个持续了多年、内容

还增加了三章的项目怀着的巨大耐心。他给了我们发挥的空间去做正确的事，我们对此铭记于心。

我还想感谢其他一些人，正是因为他们对我职业生涯的影响才有了这本书：我的父母（Lewis 和 Connie）、Andy Novobiliski 博士（研究生院）、Mina Sartipi 博士（论文导师）、Billy Harris 博士（教授研究生算法课程）、Joe Dumas 博士（研究生院）、Ritchie Carroll（openPDC 创始人）、Paul Trachian、Christophe Bisciglia 和 Mike Olson（招募我到 Cloudera）、Malcom Ramey（为我提供第一份真正的编程工作）、田纳西大学查塔努加分校，还有 Lupi 比萨店（让我在读研期间免受饥饿之苦）。

最后要感谢我的妻子 Leslie，还有我的儿子 Ethan、Griffin 和 Dane。我经常工作到很晚，有时在度假时也是如此，感谢他们对我的耐心。

Adam

我要感谢 Skymind 团队的辛勤工作，正是有他们帮忙审阅本书内容，我们才能继续安心撰写书稿。我要特别感谢 Chris，他容忍了我一边创业一边写书的疯狂想法。

DL4J 开始于 2013 与 Josh 在 MLConf 的一次邂逅，现在它已经发展成为在全世界使用的项目。DL4J 把我带向了全世界，并真正打开了我的世界，带给我无数新体验。

首先，我要感谢我的合著者 Josh Patterson，他编写了本书的大部分，大部分功劳都应该归他。这些年来，为了使本书面世，他投入了许多夜晚和周末，而我则继续研究代码库，并不断地将内容转化成新的特性。

像 Josh 一样，我们的许多队友和贡献者都尽职尽责地帮我们检查数学上是否有错误，如早期加入的 Alex、Melanie 和 Vyacheslav “Raver” Kokorin，以及后期加入的 Dave。

Tim McGovern 耐心听取了我对 O'Reilly 图书内容的一些疯狂想法，我也很感谢他让我为本书取名。

电子书

扫描如下二维码，即可购买本书电子版。



目录

前言	xv
第 1 章 机器学习回顾	1
1.1 学习的机器	1
1.1.1 机器如何学习	2
1.1.2 生物学的启发	4
1.1.3 什么是深度学习	5
1.1.4 钻进奇幻的兔子洞	5
1.2 提出问题	6
1.3 机器学习背后的数学：线性代数	7
1.3.1 标量	7
1.3.2 向量	7
1.3.3 矩阵	8
1.3.4 张量	8
1.3.5 超平面	8
1.3.6 相关数学运算	8
1.3.7 将数据转换成向量	9
1.3.8 方程组求解	10
1.4 机器学习背后的数学：统计学	12
1.4.1 概率	12
1.4.2 条件概率	14
1.4.3 后验概率	14
1.4.4 分布	15
1.4.5 样本与总体	16
1.4.6 重采样方法	16

1.4.7	选择性偏差	17
1.4.8	似然	17
1.5	机器学习如何工作	17
1.5.1	回归	17
1.5.2	分类	19
1.5.3	聚类	19
1.5.4	欠拟合与过拟合	20
1.5.5	优化	20
1.5.6	凸优化	21
1.5.7	梯度下降	22
1.5.8	SGD	24
1.5.9	拟牛顿优化方法	24
1.5.10	生成模型与判别模型	25
1.6	逻辑回归	25
1.6.1	逻辑函数	26
1.6.2	理解逻辑回归的输出	26
1.7	评估模型	27
1.8	建立对机器学习的理解	30
第 2 章	神经网络基础与深度学习	31
2.1	神经网络	31
2.1.1	生物神经元	33
2.1.2	感知器	34
2.1.3	多层前馈网络	37
2.2	训练神经网络	42
2.3	激活函数	49
2.3.1	线性函数	49
2.3.2	sigmoid 函数	49
2.3.3	tanh 函数	50
2.3.4	hard tanh 函数	51
2.3.5	softmax 函数	51
2.3.6	修正线性函数	51
2.4	损失函数	53
2.4.1	损失函数的符号	53
2.4.2	用于回归的损失函数	54
2.4.3	用于分类的损失函数	56
2.4.4	用于重建的损失函数	57
2.5	超参数	58
2.5.1	学习率	58
2.5.2	正则化	59

2.5.3	动量	59
2.5.4	稀疏	59
第 3 章	深度网络基础	60
3.1	定义深度学习	60
3.1.1	什么是深度学习	60
3.1.2	本章结构	67
3.2	深度网络的通用构建原则	67
3.2.1	参数	68
3.2.2	层	68
3.2.3	激活函数	69
3.2.4	损失函数	70
3.2.5	优化算法	71
3.2.6	超参数	73
3.2.7	小结	77
3.3	深度网络的构造块	77
3.3.1	RBM	78
3.3.2	自动编码器	82
3.3.3	变分自动编码器	83
第 4 章	深度网络的主要架构	85
4.1	UPN	85
4.1.1	DBN	86
4.1.2	GAN	88
4.2	CNN	91
4.2.1	生物学启发	92
4.2.2	思路	92
4.2.3	CNN 架构概要	93
4.2.4	输入层	94
4.2.5	卷积层	95
4.2.6	池化层	101
4.2.7	全连接层	102
4.2.8	CNN 的其他应用	102
4.2.9	CNN 列表	103
4.2.10	小结	103
4.3	RNN	103
4.3.1	时间维度建模	104
4.3.2	三维空间输入	105
4.3.3	为什么不是马尔可夫模型	107
4.3.4	常见的 RNN 架构	107

4.3.5	LSTM 网络	108
4.3.6	特定领域应用与混合网络	114
4.4	递归神经网络	115
4.4.1	网络架构	115
4.4.2	递归神经网络的变体	115
4.4.3	递归神经网络的应用	116
4.5	小结与讨论	116
4.5.1	深度学习会使其他算法过时吗	116
4.5.2	不同的问题有不同的最佳方法	117
4.5.3	什么时候需要深度学习	117
第 5 章	建立深度网络	118
5.1	将深度网络与适合的问题匹配	118
5.1.1	列式数据与多层感知器	119
5.1.2	图像与 CNN	119
5.1.3	时间序列与 RNN	120
5.1.4	使用混合网络	121
5.2	DL4J 工具套件	121
5.2.1	向量化与 DataVec	121
5.2.2	运行时与 ND4J	121
5.3	DL4J API 的基本概念	123
5.3.1	加载与保存模型	123
5.3.2	为模型获取输入	124
5.3.3	建立模型架构	124
5.3.4	训练与评估	125
5.4	使用多层感知器网络对 CSV 数据建模	126
5.4.1	建立输入数据	128
5.4.2	确定网络架构	128
5.4.3	训练模型	131
5.4.4	评估模型	131
5.5	利用 CNN 对手写图像建模	132
5.5.1	使用 LeNet CNN 的 Java 代码示例	132
5.5.2	加载及向量化输入图像	134
5.5.3	DL4J 中用于 LeNet 的网络架构	135
5.5.4	训练 CNN 网络	138
5.6	基于 RNN 的序列数据建模	139
5.6.1	通过 LSTM 生成莎士比亚风格作品	139
5.6.2	基于 LSTM 的传感器时间序列分类	146
5.7	利用自动编码器检测异常	152