

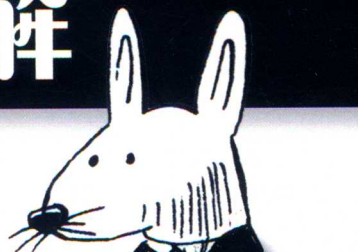
MANNING

Grokking Deep Learning

# 深度学习图解

[美] 安德鲁·特拉斯克(Andrew W. Trask) 著

王晓雷 严烈 译



清华大学出版社

# 深度学习图解

[美] 安德鲁·特拉斯克(Andrew W. Trask) 著

王晓雷 严 烈 译

贵州师范学院内部使用

清华大学出版社

北 京

Andrew W. Trask

Grokking Deep Learning

EISBN: 978-1-61729-370-2

Original English language edition published by Manning Publications, USA© 2018 by Manning Publications. Simplified Chinese-language edition copyright© 2019 by Tsinghua University Press Limited. All rights reserved.

北京市版权局著作权合同登记号 图字: 01-2019-3033

本书封面贴有清华大学出版社防伪标签, 无标签者不得销售。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13701121933

### 图书在版编目(CIP)数据

深度学习图解 / (美) 安德鲁·特拉斯克 著; 王晓雷, 严烈译. —北京: 清华大学出版社, 2020

书名原文: Grokking Deep Learning

ISBN 978-7-302-54099-1

I. ①深… II. ①安… ②王… ③严… III. ①机器学习—图解 IV. ①TP181-64

中国版本图书馆 CIP 数据核字(2019)第 241987 号

责任编辑: 王 军 韩宏志

封面设计: 孔祥峰

版式设计: 思创景点

责任校对: 牛艳敏

责任印制: 刘海龙

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社 总 机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质 量 反 馈: 010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

印 装 者: 三河市金元印装有限公司

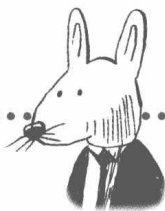
经 销: 全国新华书店

开 本: 170mm×240mm 印 张: 18.25 字 数: 330 千字

版 次: 2020 年 1 月第 1 版 印 次: 2020 年 1 月第 1 次印刷

定 价: 99.00 元

产品编号: 082781-01



## 译者序

《深度学习图解》并不是第一本有关深度学习的书，却是我所见过的最深入浅出的一本深度学习的入门读物。

在承接这本书的翻译工作之前，我已经做过数百场关于人工智能的咨询和培训。从概率统计到深度学习，从损失函数到特征工程，每一次与新的求知者接触，都会带来一轮教学相长的洗礼。我也逐渐发现，科班出身的老师往往假设听课的人已经掌握了线性代数与概率论等相关内容，知识诅咒带来的层出不穷的专业名词让读者望而却步；而专注实践的开发者往往依赖于高度封装的计算框架(比如 TensorFlow 和 PyTorch)，虽然大大降低了深度学习开发的门槛，但也在某种程度上推动着它往“玄学”方向发展。

因此，《深度学习图解》令人眼前一亮。与一般意义上的教材不同，书中大量使用了第一人称和第二人称的表达，力图以一种口语化的表达方式讲解深度学习的概念和实践。更令人惊讶的是，书中不曾出现繁杂的公式，也不曾引用先进的框架；从卷积网络到递归网络，从损失函数到反向传播，一项项复杂的概念以生动的形式展开。书中所有的代码仅基于原生 Python 和 NumPy，这更是一大亮点，读者可以亲手敲出每一行代码，检验每一点假设，一砖一瓦，为读者在深度学习领域开展进一步研究打下基础。

到 2025 年，全球人工智能市场的价值将达到约 1900 亿美元，中国已经成为全球人工智能投融资规模最大的国家。希望本书能够引领更多的人了解深度学习，掌握深度学习，更重要的是，喜欢深度学习。

也许，下一代人工智能的源泉，就在热爱创造、敬畏智能的你我之中。

王晓雷

---

## 作者简介

Andrew W. Trask 是 Digital Reasoning 公司机器学习实验室的创始成员，该实验室致力于自然语言处理、图像识别和音频转录的深度学习研究。几个月内，Andrew 和他的伙伴们就在情绪分类和词性标注方面发表了超过业界最佳方案的结果。

他训练了世界上最大的人工神经网络，拥有超过 1600 亿个参数，实验结果发表在 ICML(International Conference on Machine Learning)上，还有一部分结果发表在 Journal of Machine Learning(JML)上。他在 Digital Reasoning 公司担任文本处理和音频分析的产品经理，负责仿真认知计算平台的架构设计，深度学习是这一平台的核心能力。

.....

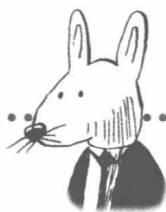
致妈妈：

你牺牲了这么多时间来教育 Tara 和我。我希望你能看到本书背后所凝结的你的心血。

也致爸爸：

谢谢你这么爱我们，在我小时候就花时间来教我学习编程技术。如果没有你，我就不会有今天的成就。

能成为你们的儿子是我的荣幸。



# 前言

这本《深度学习图解》是我三年艰苦努力的成果。为了完成你手中所握的这本书，我实际撰写的篇幅至少是你现在看到的两倍。在本书准备出版之前，有六个章节被从头重写了三到四次，同时，在这个过程中，一些计划之外的重要章节也被添加进来。

更重要的是，我很早就做出了两个决定，这两个决定使这本《深度学习图解》具有独特的价值。首先，除了基础的算术之外，这本书不需要读者具有任何数学背景，其次，它不依赖于那些可能隐藏着运行细节的高级库。换句话说，任何人都可以读懂本书，了解深度学习真正的工作方式。为达到这个目的，我必须寻找一些新的方法来描述和讲授书中的核心思想和技术，而不是依赖于别人编写的高等数学或复杂代码。

写作这本《深度学习图解》的目的是为深度学习的入门实践尽可能扫除障碍。在本书中，你不只需要阅读理论，更要靠自己发现深度学习的真谛。为了帮助你实现这一点，我编写了大量代码，并尽我所能以正确的顺序解释它——使得书中演示的每一段代码都容易理解。

深入浅出的知识，加上你将在本书中探索的所有理论、代码和示例，将使你以最快的速度完成各种实验。你可以更快地获得成功，得到更好的工作机会，甚至能够更快地掌握更先进的深度学习概念。

在过去的三年里，我不仅写了这本书，还收到了牛津大学的博士录取通知书，加入了谷歌的团队，进行了前沿研究——开发了一个去中心化的公司人工智能平台，名为 **OpenMined**。本书是我多年思考、学习和教学的结晶。

当然，深度学习还有其他很多资源。我很高兴你能选择阅读本书。

---

# 致 谢

我非常感谢每一位为《深度学习图解》的出版做出贡献的人士。首先，我要感谢 Manning 出色的团队: Bert Bates, 教我如何写作; Christina Taylor, 耐心地鼓励我坚持了三年; Michael Stephens, 他的创造力使本书在出版之前就获得了巨大的成功; 还有 Marjan Bace, 她的鼓励在我日复一日的拖延中起到了至关重要的作用。如果没有早期读者的巨大贡献, 这本《深度学习图解》就不会有今天的成就。这些读者通过电子邮件、Twitter 和 GitHub 等渠道提出意见。我非常感谢 Jascha Swisher、Varun Sudhakar、Francois Chollet、Frederico Vitorino、Cody Hammond、Mauricio Maroto Arrieta、Aleksandar Dragosavljevic、Alan Carter、Frank Hinek、Nicolas Benjamin Hocker、Hank Meisse、Wouter Hibma、Joerg Rosenkranz、Alex Vieira 和 Charlie Harrington, 感谢你们对本书文稿和在线代码库提供的帮助。

我要感谢如下这些审稿人, 他们在本书撰写期间花费宝贵时间阅读了手稿: Alexander A. Myltsev、Amit Lamba、Anand Saha、Andrew Hamor、Cristian Barrientos、Montoya、Eremey Valetov、Gerald Mack、Ian Stirk、Kalyan Reddy、Kamal Raj、Kelvin D. Meeks、Marco Paulo dos Santos Nogueira、Martin Beer、Massimo Ilario、Nancy W. Grady、Peter Hampton、Sebastian Maldonado、Shashank Gupta、Tymoteusz Wołodźko、Kumar Unnikrishnan、Vipul Gupta、Will Fuger 和 William Wheeler。

我也很感谢 Udacity 的 Mat 和 Niko, 他们把这本书选为 Udacity 的深度学习纳米学位课程的课本, 这极大地有助于本书在年轻的深度学习实践者中的推广。

还必须感谢 William Hooper 博士, 他让我走进他的办公室, 向他唠叨计算机科学的琐碎, 还破例让我进入他(已经满了)的程序设计初级班。也是他启发我从事深度学习的事业。我非常感谢他在旅程开始时对我的耐心。他对我真是太好了。

最后我要感谢我的妻子, 在我忙着写这本书的所有夜晚和周末, 她付出了极大的耐心, 与此同时, 她独立将整本书修订了好几遍, 并创建和调试了 GitHub 在线代码库。

---

# 关于本书

《深度学习图解》旨在帮助你在深度学习领域打下基础，以便能够从更高层面掌握深度学习的主要框架。它从关注神经网络的基础概念开始，然后深入讲解那些更高级的网络设计和架构。

## 目标读者

写作本书时，我有意将本书的门槛降到我认为是最底的程度。阅读本书，不需要提前掌握线性代数、微积分、凸优化甚至机器学习等任何知识。理解深度学习所需的一切知识都会在本书的阅读过程中得到解释。如果你学过高中数学，并且能够使用 Python 编程，那么已经为阅读本书做好了准备。

## 路线图

本书共有 16 章：

- 第 1 章重点介绍为什么应该掌握深度学习，以及需要做好哪些准备。
- 第 2 章开始深入探讨基本概念，比如机器学习、参数和非参数模型、监督和无监督学习。在本章中，作者首次提出了“预测、比较、学习”这一模式，该模式将在接下来的章节中进一步介绍。
- 第 3 章将带领你使用简单的网络进行预测，帮助你初步了解神经网络。
- 第 4 章将介绍如何评估第 3 章中的预测结果，识别误差，为下一步训练模型奠定基础。
- 第 5 章主要关注“预测、比较、学习”模式的“学习”部分。本章将通过一个更深入的例子来介绍“学习”的过程。
- 第 6 章中，你将学会建立第一个“深度”神经网络，包括代码和所有细节。
- 第 7 章概述神经网络，形象化你对它的理解。
- 第 8 章介绍过拟合、dropout 和批量梯度下降，并教你如何利用刚刚构建

的新网络对数据集进行分类。

- 第 9 章讲解激活函数以及如何在对概率建模时使用它们。
- 第 10 章介绍卷积神经网络，强调了卷积神经网络的结构在应对过拟合上所具有的优势。
- 第 11 章对自然语言处理(NLP)进行了深入探讨，讲解这一研究方向在深度学习领域的基础概念。
- 第 12 章讨论递归神经网络，这是近十年来在序列建模领域最先进的方法，也是业界最流行的工具之一。
- 第 13 章将让你成为深度学习框架的高级用户，帮助你从头开始快速构建深度学习框架。
- 第 14 章教你使用递归神经网络来处理一个更具挑战性的任务：语言建模。
- 第 15 章主要关注数据的隐私性，介绍基本的隐私概念，如联邦学习、同态加密以及与差分隐私和安全多方计算相关的概念。
- 第 16 章将提供继续深度学习之旅所需的其他工具和资源。

## 代码下载

本书中的所有代码都以等宽字体显示，以便将其与普通的文本区分开来。部分代码注释会伴有相应列表，以强调重要概念。

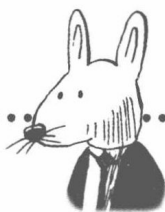
可以从出版商的网站下载本书示例的代码，网址为 [www.manning.com/books/grokking-deep-learning](http://www.manning.com/books/grokking-deep-learning) 或 <https://github.com/iamtrask/grokking-deep-learning>。

也可扫封底二维码直接下载。

## 在线论坛

购买《深度学习图解》会获赠一份权益；你可以免费访问 Manning 出版公司 (Manning Publications) 运营的一个私人网络论坛，在那里你可以对这本书发表评论，咨询技术问题，并寻求来自作者和其他用户的帮助。要访问论坛，请点击网址 <https://forums.manning.com/forums/grokking-deep-learning>。你也可以通过 <https://forums.manning.com/forums/about> 了解 Manning 论坛的行为规则。

Manning 承诺为读者提供一个有意义的交流场所，希望各位读者之间、读者与作者之间都能够进行讨论。但并未对作者的具体回应频率做出任何承诺，作者对论坛的参与是自愿的(和无偿的)。我们建议你可以试着问作者一些具有挑战性的问题，以免他失去兴趣。只要这本书还在出版期内，读者就可以从出版商的网站上看到之前的所有讨论内容。



# 目录

<b>第 1 章 深度学习简介：为什么应该学习深度学习</b> .....1	<b>第 3 章 神经网络预测导论：前向传播</b> ..... 17
1.1 欢迎阅读《深度学习图解》.....1	3.1 什么是预测..... 17
1.2 为什么要学习深度学习.....2	3.2 能够进行预测的简单神经网络..... 19
1.3 这很难学吗?.....3	3.3 什么是神经网络?..... 20
1.4 为什么要阅读本书.....3	3.4 这个神经网络做了什么?..... 21
1.5 准备工作.....4	3.5 使用多个输入进行预测..... 23
1.6 你可能需要掌握一部分 Python 知识.....5	3.6 多个输入：这个神经网络做了什么?..... 24
1.7 本章小结.....6	3.7 多个输入：完整的可运行代码..... 29
<b>第 2 章 基本概念：机器该如何学习?</b> .....7	3.8 预测多个输出..... 30
2.1 什么是深度学习?.....7	3.9 使用多个输入和输出进行预测..... 32
2.2 什么是机器学习?.....8	3.10 多输入多输出神经网络的工作原理..... 33
2.3 监督机器学习.....9	3.11 用预测结果进一步预测..... 35
2.4 无监督机器学习..... 10	3.12 NumPy 快速入门..... 37
2.5 参数学习和非参数学习..... 10	3.13 本章小结..... 40
2.6 监督参数学习..... 11	
2.7 无监督参数学习..... 13	
2.8 非参数学习..... 14	
2.9 本章小结..... 15	

## 第4章 神经网络学习导论：梯度

## 下降.....41

- 4.1 预测、比较和学习.....41
- 4.2 什么是比较.....42
- 4.3 学习.....42
- 4.4 比较：你的神经网络是否做出了好的预测？.....43
- 4.5 为什么需要测量误差？.....44
- 4.6 最简单的神经学习形式是什么？.....45
- 4.7 冷热学习.....46
- 4.8 冷热学习的特点.....47
- 4.9 基于误差调节权重.....48
- 4.10 梯度下降的一次迭代.....50
- 4.11 学习就是减少误差.....52
- 4.12 回顾学习的步骤.....54
- 4.13 权重增量到底是什么？.....55
- 4.14 狭隘的观点.....57
- 4.15 插着小棍的盒子.....58
- 4.16 导数：两种方式.....59
- 4.17 你真正需要知道的.....60
- 4.18 你不需要知道的.....60
- 4.19 如何使用导数来学习.....61
- 4.20 看起来熟悉吗？.....62
- 4.21 破坏梯度下降.....63
- 4.22 过度修正的可视化.....64
- 4.23 发散.....65
- 4.24 引入  $\alpha$ .....66
- 4.25 在代码中实现  $\alpha$ .....66
- 4.26 记忆背诵.....67

## 第5章 通用梯度下降：一次学习多个权重.....69

- 5.1 多输入梯度下降学习.....69

- 5.2 多输入梯度下降详解.....71
- 5.3 回顾学习的步骤.....75
- 5.4 单项权重冻结：它有什么作用？.....77
- 5.5 具有多个输出的梯度下降学习.....79
- 5.6 具有多个输入和输出的梯度下降.....81
- 5.7 这些权重学到了什么？.....83
- 5.8 权重可视化.....85
- 5.9 点积(加权和)可视化.....86
- 5.10 本章小结.....87

## 第6章 建立你的第一个深度神经

## 网络：反向传播.....89

- 6.1 交通信号灯问题.....89
- 6.2 准备数据.....91
- 6.3 矩阵和矩阵关系.....92
- 6.4 使用 Python 创建矩阵.....95
- 6.5 建立神经网络.....96
- 6.6 学习整个数据集.....97
- 6.7 完全、批量和随机梯度下降.....97
- 6.8 神经网络对相关性的学习.....98
- 6.9 向上与向下的压力.....99
- 6.10 边界情况：过拟合.....101
- 6.11 边界情况：压力冲突.....101
- 6.12 学习间接相关性.....103
- 6.13 创建关联.....104
- 6.14 堆叠神经网络：回顾.....105
- 6.15 反向传播：远程错误归因.....106

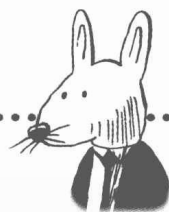
6.16	反向传播：为什么有效?.....	107
6.17	线性与非线性.....	107
6.18	为什么神经网络仍然不起作用.....	109
6.19	选择性相关的秘密.....	110
6.20	快速冲刺.....	111
6.21	你的第一个深度神经网络.....	111
6.22	反向传播的代码.....	112
6.23	反向传播的一次迭代.....	114
6.24	整合代码.....	116
6.25	为什么深度网络这么重要?.....	117
<b>第 7 章 如何描绘神经网络：在脑海里，在白纸上..... 119</b>		
7.1	到了简化的时候了.....	119
7.2	关联抽象.....	120
7.3	旧的可视化方法过于复杂.....	121
7.4	简化版可视化.....	122
7.5	进一步简化.....	123
7.6	观察神经网络是如何进行预测的.....	124
7.7	用字母而不是图片来进行可视化.....	125
7.8	连接变量.....	126
7.9	信息整合.....	127
7.10	可视化工具的重要性.....	127
<b>第 8 章 学习信号，忽略噪声：正则化和批处理介绍..... 129</b>		
8.1	用在 MNIST 上的三层网络.....	129
8.2	好吧，这很简单.....	131
8.3	记忆与泛化.....	132
8.4	神经网络中的过拟合.....	133
8.5	过拟合从何而来.....	134
8.6	最简单的正则化：提前停止.....	135
8.7	行业标准正则化：dropout.....	136
8.8	为什么 dropout 有效：整合是有效的.....	137
8.9	dropout 的代码.....	137
8.10	在 MNIST 数据集上对 dropout 进行测试.....	139
8.11	批量梯度下降.....	140
8.12	本章小结.....	143
<b>第 9 章 概率和非线性建模：激活函数..... 145</b>		
9.1	什么是激活函数?.....	145
9.2	标准隐藏层激活函数.....	148
9.3	标准输出层激活函数.....	149
9.4	核心问题：输入具有相似性.....	151
9.5	计算 softmax.....	152
9.6	激活函数使用说明.....	153
9.7	将增量与斜率相乘.....	156
9.8	将输出转换为斜率(导数).....	157
9.9	升级 MNIST 网络.....	157
<b>第 10 章 卷积神经网络概论：关于边与角的神经学习..... 161</b>		
10.1	在多个位置复用权重.....	161
10.2	卷积层.....	162

10.3	基于 NumPy 的简单实现	164	12.3	平均词向量的神奇力量	191
10.4	本章小结	167	12.4	信息是如何存储在向量嵌入中的?	192
<b>第 11 章</b>	<b>能够理解自然语言的神经网络: 国王-男人+女人=?</b>	<b>169</b>	12.5	神经网络是如何使用嵌入的?	193
11.1	理解语言究竟是指什么?	170	12.6	词袋向量的局限	194
11.2	自然语言处理(NLP)	170	12.7	用单位向量求词嵌入之和	195
11.3	监督 NLP 学习	171	12.8	不改变任何东西的矩阵	196
11.4	IMDB 电影评论数据集	172	12.9	学习转移矩阵	197
11.5	在输入数据中提取单词相关性	173	12.10	学习创建有用的句子向量	198
11.6	对影评进行预测	174	12.11	Python 下的前向传播	199
11.7	引入嵌入层	175	12.12	如何反向传播?	200
11.8	解释输出	177	12.13	让我们训练它!	201
11.9	神经网络结构	178	12.14	进行设置	201
11.10	单词嵌入表达的对比	180	12.15	任意长度的前向传播	202
11.11	神经元是什么意思?	181	12.16	任意长度的反向传播	203
11.12	完形填空	182	12.17	任意长度的权重更新	204
11.13	损失函数的意义	183	12.18	运行代码, 并分析输出	205
11.14	国王-男人+女人~=女王	186	12.19	本章小结	207
11.15	单词类比	187	<b>第 13 章</b>	<b>介绍自动优化: 搭建深度学习框架</b>	<b>209</b>
11.16	本章小结	188	13.1	深度学习框架是什么?	209
<b>第 12 章</b>	<b>像莎士比亚一样写作的神经网络: 变长数据的递归层</b>	<b>189</b>	13.2	张量介绍	210
12.1	任意长度的挑战	189			
12.2	做比较真的重要吗?	190			

13.3	自动梯度计算(autograd)		14.2	截断式反向传播的	
	介绍	211		必要性	240
13.4	快速检查	213	14.3	截断式反向传播	241
13.5	多次使用的张量	214	14.4	输出样例	244
13.6	升级 autograd 以支持多次		14.5	梯度消失与梯度激增	245
	使用的张量	215	14.6	RNN 反向传播的	
13.7	加法的反向传播如何			小例子	246
	工作?	217	14.7	长短期记忆(LSTM)	
13.8	增加取负值操作的			元胞	247
	支持	218	14.8	关于 LSTM 门限的直观	
13.9	添加更多函数的支持	219		理解	248
13.10	使用 autograd 训练神经		14.9	长短期记忆层	249
	网络	222	14.10	升级字符语言模型	250
13.11	增加自动优化	224	14.11	训练 LSTM 字符语言	
13.12	添加神经元层类型的			模型	251
	支持	225	14.12	调优 LSTM 字符语言	
13.13	包含神经元层的			模型	252
	神经元层	226	14.13	本章小结	253
13.14	损失函数层	227			
13.15	如何学习一个框架	228	<b>第 15 章</b>	<b>在看不见的数据上做深度</b>	
13.16	非线性层	228		<b>学习: 联邦学习导论</b>	255
13.17	嵌入层	230	15.1	深度学习的隐私问题	255
13.18	将下标操作添加到		15.2	联邦学习	256
	autograd	231	15.3	学习检测垃圾邮件	257
13.19	再看嵌入层	232	15.4	让我们把它联邦化	259
13.20	交叉熵层	233	15.5	深入联邦学习	260
13.21	递归神经网络层	235	15.6	安全聚合	261
13.22	本章小结	238	15.7	同态加密	262
			15.8	同态加密联邦学习	263
			15.9	本章小结	264
<b>第 14 章</b>	<b>像莎士比亚一样写作:</b>		<b>第 16 章</b>	<b>往哪里去: 简要指引</b>	265
	<b>长短期记忆网络</b>	239			
14.1	字符语言建模	239			

# 深度学习简介： | 第 1 章

## 为什么应该学习深度学习



本章主要内容：

- 为什么应该学习深度学习
- 为什么应该阅读本书
- 需要做好什么样的准备

不要担心你在数学上遇到的困难。我可以向你保证，我遇到的困难甚至更大。

——阿尔伯特·爱因斯坦

### 1.1 欢迎阅读《深度学习图解》

你将掌握 21 世纪最有价值的技能之一！

我很兴奋，你读到了这里！相信你也同样感到兴奋！深度学习代表令人热血沸腾的机器学习和人工智能的交叉领域，以及对社会和行业的重大颠覆。本书所讨论的方法正在改变你周围的世界。从汽车引擎的优化，到决定你在社交媒体上会看到哪些内容，它无处不在，无所不能，更幸运的是，它很有趣！

## 1.2 为什么要学习深度学习

它是智能增量自动化的强大工具。

从创世之初，为了更好地理解和控制我们周围的环境，人类就开始不断地发展和建造工具。在这个关于创新的叙事中，深度学习是刚刚展开的奇妙篇章。

也许，让这一章如此引人注目的是这个领域比起机械上的创新来说更像是智力上的创新。就像它在机器学习中的姐妹领域一样，深度学习试图一点一点地将智能自动化。在过去几年里，它已经在这方面取得了巨大的成功和进步，打破了历史上计算机视觉、语音识别、机器翻译等许多方面的记录。

深度学习看起来使用了一套大体上相同的受到大脑机制启发的算法(神经网络)，却在大量不同的领域取得了一系列成就——这一点尤其不同寻常。尽管深度学习仍然是一个具有诸多挑战的积极发展的领域，但它近期的发展已经让人兴奋不已：也许，我们不仅发现了一款卓越的工具，还发现了一扇通向我們自己的思维的窗户。

深度学习具有使熟练劳动力实现重大自动化的潜力。

当前，在不同的发展速度假设下，存在着大量根据当前的发展趋势来推断深度学习的潜在作用的虚假宣传。尽管这些预测中有许多过于乐观，但我相信其中有一项值得你考虑：替代人类的工作。我之所以认为这一论断值得思考，是因为即使深度学习的创新今天就停止了，它也会对遍布全球的熟练工人产生难以置信的影响。呼叫中心接线员、出租车司机和初级业务分析师都是极具说服力的例子。在这些领域，深度学习可以提供低成本的替代方案。

幸运的是，经济发展尚未出现颠覆性变化；但考虑到目前的技术力量，在很多方面我们已经要开始担心了。我希望这本书能让你(和你认识的人)从可能是那些面临威胁的行业中，迁移到这个逐渐成熟与繁荣的行业：深度学习。

它有趣又有创意。通过对智力和创造力的模拟，你会发现许多人类的秘密。

就我个人而言，我之所以踏入深度学习领域，是因为它本身就已经足够迷人。这是一个人与机器之间的令人惊奇的交叉。揭开思考、推理、甚至创造后面的秘密这件事情本身就是富有启迪的和吸引人的，足够鼓舞人心。不妨想一下，我们可以建立一个数据集，里面装满人们曾经画过的每一幅画，然后用它来教机器如何像莫奈那样画画。令人疯狂的是，这居然已经变成了现实。我们可以知道它是如何工作的——真是太棒了！