

MATLAB深度学习

机器学习、神经网络与人工智能

[美] Phil Kim

著

敖富江 杜静 周浩

译

Apress®



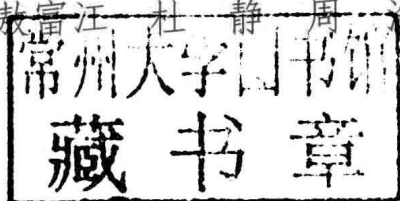
清华大学出版社

MATLAB 深度学习

机器学习、神经网络与人工智能

[美] Phil Kim 著

敖富江 杜静 周浩 译



清华大学出版社

北京

Phil Kim

MATLAB Deep Learning: with Machine Learning, Neural Networks and Artificial Intelligence

EISBN: 978-1-4842-2844-9

Original English language edition published by Apress Media. Copyright © 2017 by Apress Media. Simplified Chinese-Language edition copyright © 2018 by Tsinghua University Press. All rights reserved.

本书中文简体字版由 Apress 出版公司授权清华大学出版社出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

北京市版权局著作权合同登记号 图字：01-2017-8432

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

MATLAB 深度学习 机器学习、神经网络与人工智能 / (美)菲尔·吉米(Phil Kim)著; 敖富江, 杜静, 周浩 译.
—北京: 清华大学出版社, 2018

书名原文: MATLAB Deep Learning: with Machine Learning, Neural Networks and Artificial Intelligence

ISBN 978-7-302-49638-0

I. ①M… II. ①菲… ②敖… ③杜… ④周… III. ①Matlab 软件 IV. ①TP317

中国版本图书馆 CIP 数据核字(2018)第 033888 号

责任编辑: 王 军 韩宏志

封面设计: 牛艳敏

版式设计: 思创景点

责任校对: 曹 阳

责任印制: 王静怡

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

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 三河市国英印务有限公司

经 销: 全国新华书店

开 本: 170mm×240mm 印 张: 8.25 字 数: 152 千字

版 次: 2018 年 3 月第 1 版 印 次: 2018 年 3 月第 1 次印刷

印 数: 1~3500

定 价: 49.80 元

产品编号: 076420-01

译者序

继云计算、物联网、大数据之后，备受追捧的人工智能将在人类未来的生活中扮演越来越重要的角色。人工智能并非新兴学科，与其相关的研究与开发已有60多年。经历了在学术界与商业界几次长短不一的热潮与寒潮之后，此次引领人工智能成为热点的是已在图像识别、语音识别、自然语言处理等领域取得辉煌成就的“深度学习”技术。这一技术解决了机器学习在处理实际问题过程中曾遇到的诸多难题，为人工智能在新时期的应用与推广打下了基础。掌握深度学习这项技能对于信息化时代海量数据的自动化处理至关重要。它势必成为IT从业人员的必备技能之一。

本书是一本关于深度学习的入门书籍。作者 Phil Kim 博士是一位研究人工智能数据处理及机器学习的技术专家，同时具有丰富的 MATLAB 编程及应用经验。这使得本书既有理论深度，又紧密结合实际。本书不仅适用于在深度学习领域开展研究的学术人员，也适用于深度学习应用的开发者。本书深入浅出地介绍人工智能、机器学习、神经网络、深度学习这些关键概念以及它们之间的关联，并着重介绍监督学习这一机器学习方式。为便于读者理解，本书提供了大量示例，并利用 MATLAB 语言完成了代码实现。结合对代码的分析与讲解，本书致力于使读者更清晰地理解书中所涉及的理论知识。

本书主要由敖富江、杜静、周浩翻译，参与本书翻译的还有李海莉、张民磊、李博、鲁志勇、岁赛、庞训龙、孔德强、秦富童、袁学军等。为精确地表达原文含义，并使译文通顺、完美，译者们在翻译过程中查阅、参考了大量中英文资料，并进行了广泛讨论。当然，限于水平和精力有限，翻译中的错误和不当之处在所难免，我们非常希望能得到读者的积极反馈以利于更正和改进。

感谢本书的作者，于字里行间感受到作者的职业精神和专业素养；感谢清华大学出版社给予我们翻译本书并进行学习的机会；感谢清华大学出版社的编辑们，他们为本书的翻译校对投入了巨大的热情并付出了很多心血，没有他们的帮

助和鼓励，本书不可能顺利付梓。

最后，希望读者通过阅读本书能早日掌握深度学习的精髓，以便进一步探索人工智能与机器学习这一广袤领域！

作者简介

Phil Kim 博士是一名经验丰富的 MATLAB 程序员和用户。他致力于研究来源于人工智能的海量数据的处理算法，并且研究机器学习。他曾任韩国航天航空研究院高级研究员。在该院他的主要任务是开发用于无人驾驶飞行器的自主飞行算法和机载软件。他在攻读博士学位期间开发了一款名为 **Clickey** 的屏幕键盘程序，该程序成为他当前就任韩国国家康复研究院高级研究员的桥梁。

技术审校者简介

Jonah Lissner 是一位研究科学家，致力于推进 PhD(哲学博士)和 DSc(理学博士)人才培养、奖学金授予、应用项目，也负责理论物理、电力工程、复杂系统、超级材料、地球物理学和计算理论领域的学术期刊出版。他对用于假设建构、理论学习、抽象问题求解的数学和公理建模与测试方面的经验主义和科学推理具有很强的认知能力。有关他的论文、研究出版物、项目、简历、期刊、博客、专著和系统，请参见 <http://Lissnerresearch.weebly.com>。

致 谢

虽然我认为大多数书籍的致谢对于读者来说无关紧要，但我仍想用只言片语来表达我的感激之情，因为下面将要提到的人对我来说十分重要。首先，非常感谢在 Modulabs(www.modulabs.co.kr)与我一起研究深度学习的学者，感谢他们讲授了深度学习知识。另外，衷心感谢 Modulabs 的主任 S. Kim，他允许我在这么优雅的地方工作并度过了春天与夏天。我有幸在 Modulabs 完成了本书大部分内容的写作。

还要感谢 Bogonet 的主席 Jeon，感谢 KARI 的 H. You 博士、Y. S. Kang 博士、J. H. Lee 先生，感谢 Modulabs 的主席 S. Kim，以及 J.MARPLE 的 W. Lee 先生和 S. Hwang 先生。他们耗费了大量时间和精力阅读并修订了本书的初稿。虽然他们提出了很多棘手问题，导致我在整个修订过程中十分辛苦，但促使我无憾地完成了本书的最终撰写。

最后，将最深切的感谢和爱献给我的妻子，她是我见过的最优秀女性，同时也献给我的孩子们，他们从未厌烦我，我们一起分享了美好时光。

前 言

我很幸运地亲眼见证了世界向信息社会的转变，紧随而来的是网络化环境。自从青年时代起，我就生活在一个充满变革的时代。个人电脑打开了通向信息世界的大门，接踵而至的是通过互联网将电脑连在一起的在线通信，以及将人们联系在一起的智能手机。现在，每个人都感受到了人工智能浪潮的兴起。越来越多的智能服务正在被引入，开启一个新时代。深度学习是引领该智慧浪潮的技术。尽管它最终可能将其王位权杖移交给其他技术，但现在，它是人工智能这项新技术的基石。

深度学习非常受欢迎，几乎可在任何地方找到与之相关的资料。然而，这些资料中很少是为初学者准备的。我写这本书的目的就是希望读者可在这个领域开展研究，并且避免我在初次接触深度学习时所遭遇的困境。我也希望本书步进式的叙述方法可以帮助读者避免我以前所面对过的困惑。

本书适用于两类读者。第一类读者是那些计划系统地学习深度学习以便进一步开展研究和开发的人员。这类读者应当从头到尾阅读本书的所有内容。示例代码对于深入理解概念来说特别有帮助。本书作者着力构建了丰富的示例并对其进行实现。代码示例被设计得易于阅读和理解，它们是采用 MATLAB 编写的，以便具有更好的可阅读性。没有哪种编程语言能比 MATLAB 更好地以简单直观的方式处理深度学习中出现的问题。示例代码仅使用了一些基本函数和语法，以方便那些不熟悉 MATLAB 的读者轻松地理解这些概念。对于那些熟悉编程的人士而言，示例代码可能比本书的文字更容易理解。

另一类读者是那些想获得更深入的深度学习信息(而非杂志或报纸上的泛泛之谈)但又不想系统地研究深度学习的人员。这类读者可以跳过示例代码，简单地阅读一下概念解释。这类读者可能尤其需要略过神经网络的学习规则这一部分。实际上，由于存在多种可利用的深度学习库，即使是开发人员也很少需要实现学习规则。因此，那些从来不需要开发学习规则的读者不需要在其上费心费力。然

而，这类读者需要仔细学习第 1 章、第 2 章、第 5 章和第 6 章。第 6 章特别有助于读者掌握深度学习的最重要技术，即使是仅阅读了其中的概念和示例结果，你也将受益匪浅。书中偶尔会出现一些提供理论背景知识的公式。然而，它们只是是一些基本操作。读者仅需竭尽所能地进行阅读和学习，最终便可以全面理解这些概念。

本书的组织结构

本书共有 6 章，分为三个主题。第一个主题是机器学习，第 1 章对其进行了介绍。深度学习源于机器学习。这意味着如果读者想理解深度学习的本质，必须在一定程度上了解机器学习背后的哲学思想。第 1 章从介绍机器学习与深度学习的关系开始，接着介绍机器学习的问题求解策略和根本局限。该章不介绍具体技术。然而，该章将涵盖一些适用于神经网络和深度学习的基本概念。

第二个主题是人工神经网络¹。第 2~第 4 章重点介绍这一主题。由于深度学习是一种应用神经网络的机器学习方法，神经网络与深度学习密不可分。第 2 章介绍神经网络的基本原理开始，包括其操作原理、架构和学习规则。该章还叙述了简单的单层架构演变成复杂的多层架构的原因。第 3 章介绍反向传播算法，它既是神经网络的重要而有代表性的学习规则，也是深度学习中所使用的算法。该章解释了代价函数和学习规则相关联的方式，以及那些在深度学习中得到广泛应用的代价函数。

第 4 章介绍如何将神经网络应用于分类问题。本书为分类问题单独分配了一章是因为它是当前最流行的机器学习应用。例如，图像识别这一主要的深度学习应用便是分类问题。

第三个主题是深度学习。这也是本书的主题。第 5 章和第 6 章介绍深度学习。第 5 章介绍使深度学习发挥卓越性能的驱动因素。为便于读者理解，该章从深度学习遭遇(并解决)自身难题的发展史开始介绍。第 6 章介绍卷积神经网络，它是典型的深度学习技术。卷积神经网络是图像识别技术的不二选择。该章首先介绍卷积神经网络的基本概念和架构，并将其与之前的图像识别算法进行比较。接着介绍作为卷积神经网络基本组成部分的卷积层和池化层的作用和原理。该章最终以一个应用了卷积神经网络的数字图像识别示例结束，并分析了图像经由卷积神经网络各层时的演变情况。

1 除了可能与人脑的神经网络相混淆的情形，本书将“人工神经网络”简称为“神经网络”。

源代码

本书中使用的所有源代码均可通过 Apress 网站(www.apress.com/9781484228449)在线获得,也可扫描本书封底的二维码下载。这些示例代码已经在 MATLAB 2014a 环境中进行了测试,不再需要其他软件工具。

目 录

第 1 章 机器学习	1
1.1 什么是机器学习	2
1.2 机器学习面临的挑战	4
1.2.1 过拟合	5
1.2.2 克服过拟合	7
1.3 机器学习的类型	9
1.4 本章小结	13
第 2 章 神经网络	15
2.1 神经网络的节点	15
2.2 神经网络的层	17
2.3 神经网络的监督学习	21
2.4 训练单层神经网络: delta 规则	22
2.5 广义 delta 规则	25
2.6 SGD、Batch 和 Mini Batch	26
2.6.1 SGD	26
2.6.2 Batch	27
2.6.3 Mini Batch	27
2.7 delta 规则示例	29
2.8 SGD 方法的实现	30
2.9 Batch 方法的实现	32
2.10 SGD 与 Batch 的比较	34
2.11 单层神经网络的局限性	36

2.12	究竟发生了什么?	38
2.13	本章小结	40
第3章	多层神经网络的训练	41
3.1	反向传播算法	42
3.2	反向传播示例	46
3.2.1	XOR 问题	48
3.2.2	动量	50
3.3	代价函数与学习规则	53
3.4	交叉熵函数示例	57
3.5	交叉熵函数	58
3.6	代价函数比较	60
3.7	本章小结	62
第4章	神经网络与分类问题	63
4.1	二元分类	63
4.2	多元分类	66
4.3	多元分类示例	71
4.4	本章小结	78
第5章	深度学习	79
5.1	深度神经网络的改进	80
5.1.1	梯度消失	81
5.1.2	过拟合	82
5.1.3	计算负载	83
5.2	ReLU 与 Dropout 的实例	84
5.2.1	ReLU 函数	85
5.2.2	Dropout	88
5.3	本章小结	93
第6章	卷积神经网络	95
6.1	卷积神经网络架构	95
6.2	卷积层	97
6.3	池化层	101
6.4	MNIST 示例	102
6.5	本章小结	116

第 1 章



机器学习

很容易见到媒体将机器学习与深度学习这两个概念互换使用。然而，学术专家通常会将它们区别使用。如果读者希望学习这一领域，理解这两个名词的真正含义就很重要，而更重要则是理解它们有何差别。

初闻“机器学习”这个名词，你脑海里出现的是什么呢？是不是一个类似图 1-1 的场景？那就未免严重望文生义了。

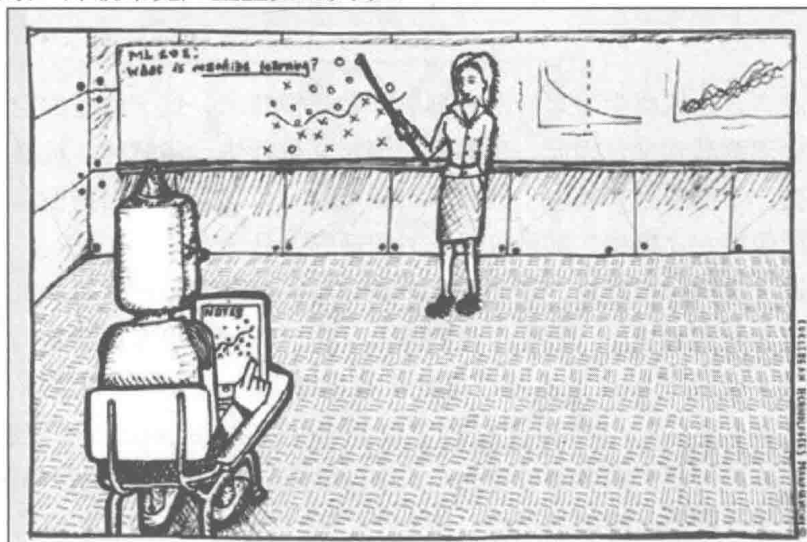


图 1-1 机器学习还是人工智能？

图 1-1 更多地描述了人工智能，而非机器学习。按此图理解机器学习的概念会导致严重的混淆。虽然机器学习确实是人工智能的一个分支，但机器学习包含的理念与该图所体现的有着极大不同。

总的来说，人工智能、机器学习和深度学习存在如下关系：

“深度学习是机器学习的一个子类，而机器学习又是人工智能的一个子类。”

这样概括怎样？是不是很简单明了？该分类也许没有自然规律那样绝对，但它是被广泛认可的。

进一步讲，人工智能是一个十分宽泛的概念，它可以描述很多不同的事物。它可以指代任何形式的蕴含某些智能特性的技术，并非特指某一特定技术领域。相反，机器学习则指一个特定领域。也就是说，机器学习这一概念用于指代人工智能的一个特定技术类别。机器学习本身也包含多种技术，其中之一就是深度学习，这也是本书要探讨的内容。

理解深度学习是机器学习的一个子类是非常关键的，这也是在此如此长篇大论探讨人工智能、机器学习与深度学习的关系的原因。由于深度学习高效地解决了一些曾制约人工智能的问题，它最近成为研究热点之一。在很多领域，它的性能都出类拔萃。尽管如此，深度学习也存在着局限。它的局限源于其基本理念是从它的父类(即机器学习)那里继承而来。作为机器学习的一个子类，深度学习无法避免机器学习所存在的根本性问题。因此，在讨论深度学习之前，本书首先回顾机器学习的基本概念。

1.1 什么是机器学习

简而言之，机器学习是针对数据的一种建模技术。这一简短的定义可能很难让初学者理解其含义。因此，下面将对该定义进行进一步解释。机器学习是一种从“数据”中抽象出“模型”的技术。顾名思义，这里的数据指信息，如文档、音频、图像等。“模型”则是机器学习的最终产品。

在进一步探讨模型这一概念前，不妨先暂时引申到另一话题。机器学习的概念中仅包含数据和模型，却完全没有涉及“学习”，这很奇怪不是吗？“学习”这一名词本身反映了该技术能自主分析数据并发现模型而不需要人工参与的特性。称之为“学习”，是因为其过程类似于利用数据进行训练以实现建模。因此，机器学习建模过程中的数据被称为“训练”数据。机器学习的流程如图 1-2 所示。

现在回到之前对模型的讨论。事实上，模型不过是我们想获得的最终产品。举个例子，如果要开发一套删除垃圾邮件的自动过滤系统，那么垃圾邮件的过滤器就是所讨论的模型。从这个角度看，可以说模型就是用户实际使用的工具。也有人把模型称为假设(hypothesis)，对具有统计学背景的读者而言该术语似乎更为易懂。

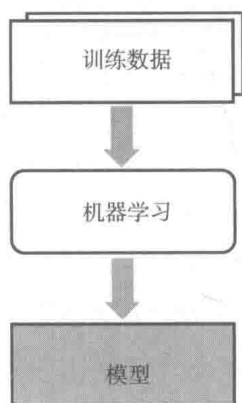


图 1-2 机器学习过程

机器学习并非唯一的建模技术。多年来在动力学领域，人们使用称为运动方程的一系列基于牛顿定律、描述物体运动规律的公式进行建模。而在人工智能领域则有专家系统，它是基于专家的知识与诀窍所构建的问题求解模型，其工作方式类似于专家本人。

遗憾的是，在某些领域，定律及逻辑推理对建模的用处并不大。其中的典型问题均涉及真正的智能，例如图像识别、语音识别和自然语言处理。在此举一个例子：请查看图 1-3，并尝试识别其中的数字。

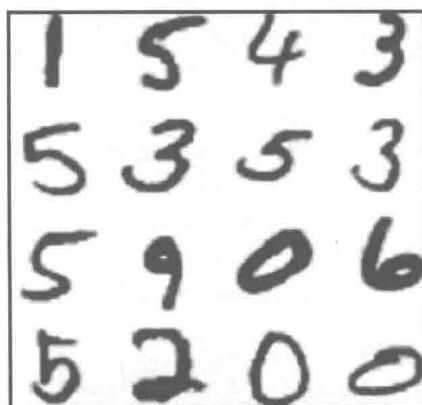


图 1-3 计算机如何辨别这些不具备可识别模式的数字？

相信读者瞬间就能完成这个任务，大部分人都可以。那么，要让计算机试着做同样的事情，应该怎么做呢？若想利用传统建模技术，就需要找到相应的规则或算法来区分离出的这些数字。不妨就利用人类大脑中辨认数字时所使用的规则吧？很简单，不是吗？其实不然。事实上，这是一个很具挑战性的问题。研究人员也曾经认为，计算机完成这个任务应该是小菜一碟，因为对于人类来说完成这个任务极其容易，而计算机具备比人类更强的计算能力。但没过多久，研究人员便意识到他们错了。

人类是如何在没有一种清晰规范或规则的情况下识别数字的？这是个很难回答的问题，不是吗？但到底是为什么呢？因为我们从未专门学习这样的规范。我们在幼年时，便直接学到了“这个是 0，那个是 1”。我们只认为事实就是如此，并在接触更多的数字时，逐渐增强数字辨识能力。是这个道理，对吧？

那么对于计算机又如何？难道不能让计算机做同样的事吗？正是这样！恭喜！这就是机器学习的理念。机器学习的诞生就是为了解决那些难以得出解析模型的问题。该技术的核心思想就是在无法使用公式及定理得到满意结果时，利用训练数据来建立模型。

1.2 机器学习面临的挑战

上一节说明了机器学习是从数据中发现(或学习)模型的技术。它适于解决涉及智能的问题，例如图像和语音识别。这些问题很难通过使用物理定律或数学公式实现建模。一方面，机器学习所采用的方法是其流程有效性的保障；另一方面，它也带来了难以避免的问题。这一节将讨论机器学习面临的根本性问题。

通过机器学习流程从训练数据中获取模型后，我们就将该模型应用于处理实际的领域数据。其过程如图 1-4 所示。图中垂直方向的流程展示了学习过程，而训练所得到模型的应用则由水平方向的流程表示，即推理。

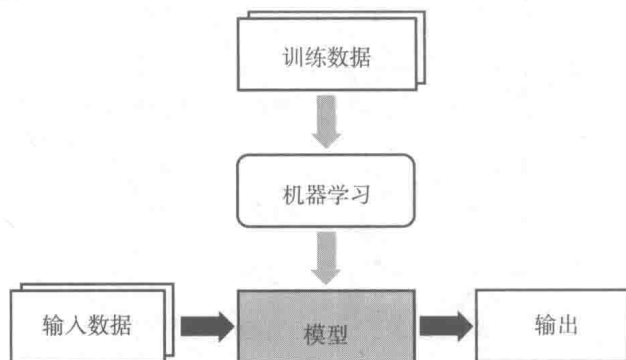


图 1-4 基于实际领域数据的应用模型

机器学习在建模过程中所使用的数据与实际应用所使用的数据是不同的。通过在图 1-4 中增加另一个模块，图 1-5 可以更好地描绘这一情况。

训练数据与输入数据之间存在差异是机器学习面临的结构性挑战。毫不夸张地说，它是机器学习所存在的一切问题的根源。比如，利用某人的手写笔迹作为训练数据，会导致何种结果？该模型能否成功识别出另一个人的笔迹？这种可能性非常低。