

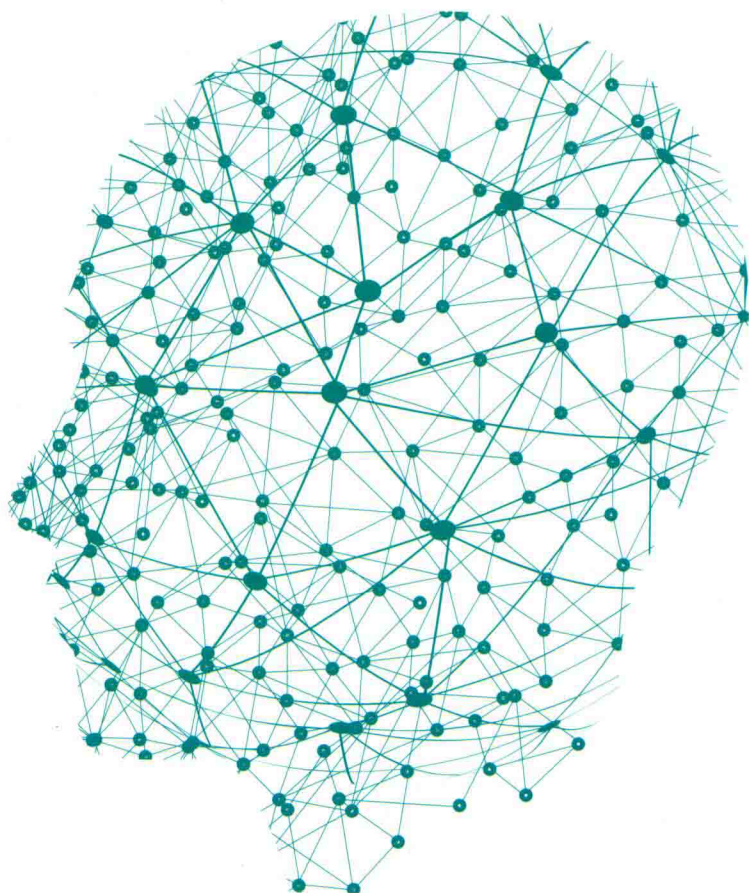
深度学习：

基于MATLAB的设计实例

Deep Learning for Beginners:
with MATLAB Examples

[韩]Phil Kim 著

邹伟 王振波 王燕妮 译



北京航空航天大学出版社
BEIHANG UNIVERSITY PRESS

深度学习： 基于 MATLAB 的设计实例

Deep Learning for Beginners:
with MATLAB Examples

[韩]Phil Kim 著

邹 伟 王振波 王燕妮 译

北京航空航天大学出版社

图书在版编目(CIP)数据

深度学习：基于 MATLAB 的设计实例 / (韩) 金晟箭
(Phil Kim) 著；邹伟，王振波，王燕妮译. -- 北京：
北京航空航天大学出版社，2018.3

书名原文：Deep Learning for Beginners: with
MATLAB Examples

ISBN 978-7-5124-2666-5

I. ①深… II. ①金… ②邹… ③王… ④王… III.
①机器学习 IV. ①TP181

中国版本图书馆 CIP 数据核字(2018)第 050355 号

英文原名: Deep Learning for Beginners: with MATLAB Examples

Copyright© 2016 by Phil Kim.

Translation Copyright© 2017 by BeiHang University Press. All rights reserved.

本书中文简体字版由 Phil Kim 授权北京航空航天大学出版社在中华人民共和国境内独家出版发行。版权所有。

北京市版权局著作权合同登记号 图字:01-2017-5532 号

深度学习：基于 MATLAB 的设计实例

Deep Learning for Beginners: with MATLAB Examples

[韩]Phil Kim 著

邹伟 王振波 王燕妮 译

责任编辑 宋淑娟

*

北京航空航天大学出版社出版发行

北京市海淀区学院路 37 号(邮编 100191) <http://www.buaapress.com.cn>

发行部电话:(010)82317024 传真:(010)82328026

读者信箱: emsbook@buaacm.com.cn 邮购电话:(010)82316936

涿州市新华印刷有限公司印装 各地书店经销

*

开本:710×1000 1/16 印张:12.5 字数:173千字

2018年4月第1版 2018年4月第1次印刷 印数:3000册

ISBN 978-7-5124-2666-5 定价:59.00元

若本书有倒页、脱页、缺页等印装质量问题,请与本社发行部联系调换。联系电话:(010)82317024

序 言

我有幸见证了世界向信息化社会的转变过程,随之而来的就是一个网络化的环境。我从小就生活在这种变革中。个人计算机的发明打开了人类通向信息世界的大门,接着就是互联网将计算机连接了起来,智能手机将人与人联系了起来。现在,每个人都意识到人工智能的浪潮已经到来。越来越多的智能服务即将被发明出来,同时这也将把我们带入一个新的智能时代。深度学习是引领这股智能浪潮的前沿技术。虽然它最终可能将其宝座移交给其他新技术,但是目前它仍是各种人工智能新技术的重要基石。

深度学习如此流行,以至于关于它的资料随处可见。然而适用于初学者的资料并不多见。我编写这本书的目的是希望帮助初学者在学习这个新知识的过程中不那么痛苦,因为我曾体验过这种痛苦,同时也希望本书中具体的开发实例讲解能够帮助初学者避免我曾经遇到的困惑。

本书主要考虑了两类读者。第一类是准备系统地学习深度学习以进一步研究和开发的读者。这类读者需要从头到尾阅读本书内容,其中的示例代码将更加有助于进一步理解本书所讲的内容。我为提供恰当的例子并加以实现做出了很大的努力,同时为了使编写的代码易于阅读和理解,均将它们用 MATLAB 编写而成。在简单和直观性上,没有任何语言比 MATLAB 更易于处理深度学习中的矩阵。示例代码仅采用了基本的函数和语法,以便不熟悉 MATLAB 的读者也能容易理解和分析里面的概念。对于熟悉编程的读者来说,代码可能比文字更容易理解。

第二类是想比从杂志或报纸上获得更深入的深度学习信息,但不必进行实际研究的读者。这类读者可以跳过代码,只需简要地阅读对这些概念

的解释即可;也可以跳过神经网络的学习规则这部分内容。实际上,因为很容易获取各种深度学习库,甚至开发者很少需要亲自实现这些学习规则,因此,对于那些从不想开发深度学习的人员,不必担心本书内容的难度。但是请重点关注第 1 章、第 2 章(第 2.1~2.4 节)、第 5 章和第 6 章的内容。特别是第 6 章,即使只是阅读其概念和示例结果,也有助于理解深度学习的大多数重要技术。为了提供理论背景,本书中偶尔会出现一些方程,但它们只是基础的运算。阅读和学习你能忍受的内容最终将让你对这些概念有个全面的理解。

本书结构

本书共包含 6 章内容,可以分为 3 个主题。

第 1 个主题是机器学习,这是第 1 章的内容。深度学习起源于机器学习,这意味着如果想要理解深度学习的本质,就必须在某种程度上知道机器学习背后的理念。第 1 章从机器学习与深度学习的关系开始讲起,随后是解决问题的策略和机器学习的基本局限性。此处仅涵盖了神经网络和深度学习的基本概念,并没有详细介绍技术本身。

第 2 个主题是人工神经网络^①,这是第 2~4 章的重点内容。由于深度学习就是采用一种神经网络的机器学习,所以不能将神经网络与深度学习分开。第 2 章从神经网络的基本概念讲起:它的工作原理、体系结构和学习规则,也讲到了神经网络由简单的单层结构演化为复杂的多层结构的原因。第 3 章介绍了反向传播算法,它是神经网络中一种重要和典型的学习规则,深度学习也使用这种算法。本章解释了代价函数和学习规则是如何联系起来的,哪一种代价函数在深度学习中被广泛使用。第 4 章介绍了将神经网络应用到分类问题中的方法。其中单列一节专门讲分类,因为它是目前最流行的一种深度学习应用。例如图像识别是一个分类问题,也是深度学习的

^① 除非它与人脑神经网络相混淆时才加以说明,本书中的神经网络指的就是人工神经网络。

一种主要应用。

第3个主题是深度学习,也是本书的重点,将在第5章和第6章中讲解。第5章介绍了使深度学习能够产生卓越性能的驱动因素。为了有助于更好地理解,本章先谈了深度学习的发展历程,包括它遇到的障碍及解决办法。第6章讲解了卷积神经网络,它是深度学习的代表性技术。卷积神经网络在图像识别领域是首屈一指的技术。本章首先介绍了卷积神经网络的基本概念和结构,并与前面的图像识别算法进行了比较;随后解释了卷积层和池化层的作用和运算方法,它们是卷积神经网络的重要组成部分。第6章也包含了一个用卷积神经网络进行数字图像识别的例子,并研究了图像通过各层的演化过程。

示例代码

本书中的全部代码和数据都能通过下面的链接获取,这些例子都通过了 MATLAB 2014 的测试,并且不需要额外的工具箱,链接地址是

github.com/philbooks/Deep-Learning-for-Beginners

致 谢

实际上,我认为大部分书籍的致谢都与读者无关,然而我还是准备按惯例写下一些感谢的话语,因为许多人和事对我来说都很特别。首先,我对在 Modulabs(www.modulabs.co.kr)共同学习深度学习的朋友们深表感谢,我所知道的大部分深度学习知识都来源于他们;并且我还要感谢我的导师 S. Kim,是他接受我,把我领进了这个奇妙的领域,共度春夏秋冬,我才能在 Modulabs 完成本书的大部分内容。

同时我也感谢来自 Bogonet 的 Jeon 主席,来自 KARI 的 H. You 博士、Y. S. Kang 博士和 J. H. Lee 先生,来自 Modulabs 的 S. Kim 导师,来自 J. MARPLE 的 W. Lee 先生和 S. Hwang 先生,他们都花了很多时间和精力

力来阅读和修改我的书稿,他们在整个修改过程中提出了很多建议,虽然这也给了我一段艰难的时光,但只有这样我现在才能毫无遗憾地完成本书。

最后,我把最深的感谢和爱献给我的妻子,她是我此生遇到的最好的女人;我也爱我的孩子,他们从来不会厌倦我,并且与我分享了珍贵的回忆。

Phil Kim

目 录

第 1 章 机器学习 /1

- 1.1 机器学习与深度学习 /3
- 1.2 什么是机器学习 /5
- 1.3 机器学习的挑战 /7
- 1.4 过拟合 /9
- 1.5 直面过拟合 /11
- 1.6 机器学习的类型 /14
- 1.7 分类和回归 /16
- 1.8 总 结 /18

第 2 章 神经网络 /21

- 2.1 概 述 /23
- 2.2 神经网络节点 /24
- 2.3 多层神经网络 /26
- 2.4 神经网络的监督学习 /33
- 2.5 单层神经网络训练:增量规则 /34
- 2.6 广义增量规则 /37
- 2.7 随机梯度下降算法、批量算法和小批量算法 /39
 - 2.7.1 随机梯度下降算法 /39

- 2.7.2 批量算法 /40
- 2.7.3 小批量算法 /41
- 2.8 示例:增量规则 /42
 - 2.8.1 随机梯度下降算法的实现 /44
 - 2.8.2 批量算法的实现 /47
 - 2.8.3 随机梯度下降算法与批量算法的比较 /51
- 2.9 单层神经网络的局限性 /54
- 2.10 总 结 /58

第 3 章 训练多层神经网络 /61

- 3.1 概 述 /63
- 3.2 反向传播算法 /65
- 3.3 示 例 /70
 - 3.3.1 XOR 问题 /72
 - 3.3.2 动量法(Momentum) /76
- 3.4 代价函数和学习规则 /80
- 3.5 示 例 /86
 - 3.5.1 交叉熵函数 /86
 - 3.5.2 代价函数的比较 /90
- 3.6 总 结 /93

第 4 章 神经网络及其分类 /95

- 4.1 概 述 /97
- 4.2 二分类 /98
- 4.3 多分类 /101
- 4.4 示例:多分类 /106

4.5 总 结 /118

第 5 章 深度学习 /119

5.1 概 述 /121

5.2 深度神经网络的进化 /123

5.2.1 梯度消失 /123

5.2.2 过拟合 /125

5.2.3 计算量的增加 /127

5.3 示 例 /128

5.3.1 ReLU 函数 /128

5.3.2 节点丢弃 /135

5.4 总 结 /143

第 6 章 卷积神经网络 /145

6.1 概 述 /147

6.2 卷积神经网络的架构 /148

6.3 卷积层 /150

6.4 池化层 /156

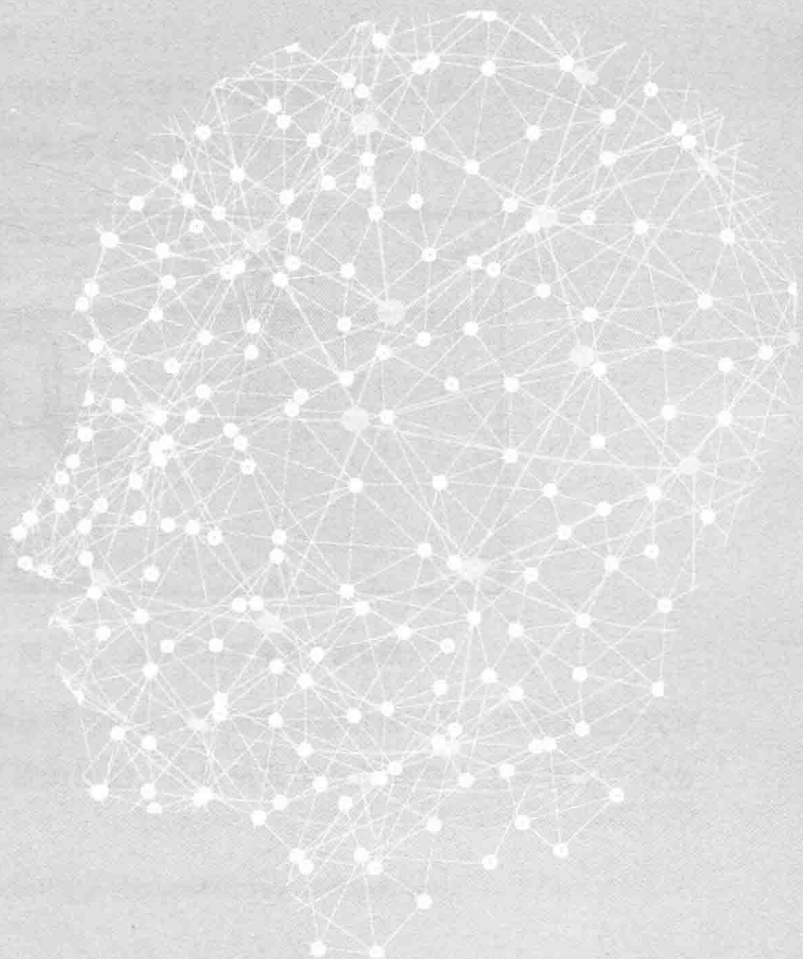
6.5 示例:MNIST /158

6.6 总 结 /178

索 引 /181

第 1 章

机器学习



1.1 机器学习与深度学习

在互联网上很容易找到机器学习和深度学习的概念,但是人们经常将它们混淆。如果已决定了这个研究方向,那么就应该分清这些词的实际含义,尤其是它们的区别。

当你第一次听到“机器学习”这个词的时候想到了什么?会不会是图 1-1 中的这个场景^①? 如果是,那么你就是只看字面意思了。

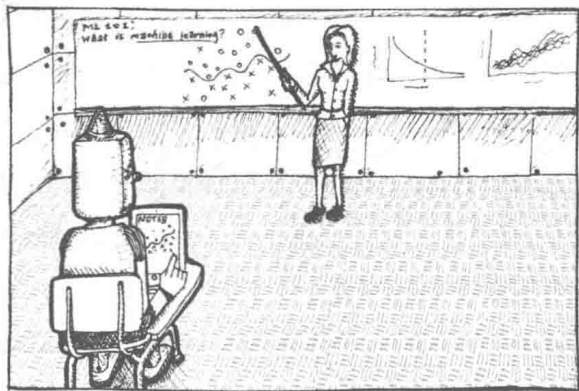


图 1-1 这是不是机器学习的场景?

图 1-1 展示的内容更偏向于人工智能而非机器学习。以图中的方式理解机器学习会带来严重的困惑。虽然机器学习的确是人工智能的一个分

^① 摘自:Euclidean Technologies Management(www.euclidean.com)。

支,但它所表达的意思与图示中的内容完全不同。

通常,人工智能、机器学习、深度学习的关系是:“深度学习是机器学习的分支,机器学习是人工智能的分支。”三者的关系可用图 1-2 来表示。

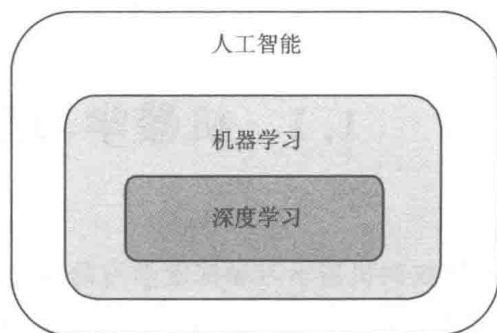


图 1-2 人工智能、机器学习、深度学习的关系

哇! 什么意思? 很简单,对不对? 这种认识可能不会像自然法则那样是绝对正确的,但是已经被广泛接受了。

下面来深挖一下这个概念。人工智能是一个范围极广的词,它可以表达很多不同的事物。它可能指任何包含一些智能的技术,而非指某些特定的技术领域。与之相反,机器学习恰恰指某些特定的领域。换句话说,使用“机器学习”这个词来特指人工智能中某些特定的领域。当然,机器学习本身也包括很多技术,其中之一就是深度学习,即本书的主题。

深度学习是机器学习的分支这一事实非常重要,这也是为什么这里长篇大论地讲人工智能、机器学习、深度学习之间的关系。深度学习最近被广泛关注,因为它高效地解决了以往对人工智能极具挑战性的任务。深度学习的效果在很多领域都表现得非常好。当然,深度学习也有限制,深度学习的一个束缚就是它源自机器学习的一些概念,作为机器学习的分支,深度学习不可避免地涉及了机器学习的各个方面。这就是为什么本书需要在深度学习之前回顾机器学习。

1.2 什么是机器学习

简言之,机器学习是通过数据进行建模的技术。这个定义或许对初学者来说太过简短了,很难真正地理解。因此,下面略微深入地解释它。机器学习就是从给定的数据中挖掘出合适模型的技术。这里的数据指的是文档、声音、图像等各类信息,模型是机器学习的最终输出结果。

在继续解释模型之前,先插入一个小话题。在机器学习的定义中只是提出了数据和模型的概念,定义中似乎与“学习”没有什么关系,这难道不奇怪吗?其实这个名字反映了分析数据和找到模型的技术方法:机器学习是依赖模型自身获得参数而非依赖人。之所以叫它“学习”是因为这个过程类似于训练这批数据去找到模型,从而解决问题。因此,机器学习在建立模型过程中所使用的数据叫做“训练数据”,图 1-3 展示了机器学习的流程。

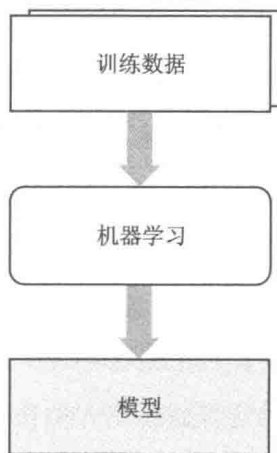


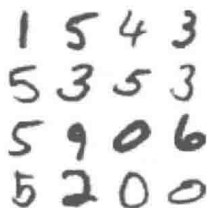
图 1-3 建立机器学习模型的流程

下面继续讨论模型。事实上,模型的唯一目的就是获得最终的结果。

例如,如果想开发一个自动滤除垃圾邮件的自动过滤系统,那么“垃圾邮件过滤器”就是需要讨论的模型。在这种情况下,就可以说实际使用的是模型。有人把模型称为“假设”,这个术语更多地表现了统计学的背景。

机器学习并非建模的唯一技术。在动力学领域中,人们早已长期使用牛顿定律和一系列动量方程来描述物体的动量;在人工智能领域中,人们使用基于知识表达和知识使用的专家系统来建模。从实践中发现,这个模型就像专家那样,效果非常好。

但是,在有些领域中却很难通过客观规律和逻辑推理来建模,典型的就是那些涉及智能领域的问题,如图像识别、语音识别和自然语言处理。比如,图 1-4 中有关数字识别的问题。



1 5 4 3
5 3 5 3
5 9 0 6
5 2 0 0

图 1-4 简单示例:手写数字图像识别

相信大家一定是在瞬间就完成了数字识别任务,而且大多数人都是如此。不过,如果让计算机做相同的事情,那么该如何设计算法模型呢?如果使用传统的建模技术,就需要找到一些规则或算法来判别手写的数字。既然如此,直接使用大脑中用到的识别数字的那些规则不就好了么?非常简单,对吧?但事实并非如此。事实上,这是一个非常有挑战性的工作。有一个时期,研究者认为这对于计算机来说是小菜一碟,因为数字识别对于人类来说是小菜一碟,而计算机的运算速度比人类快很多,所以自然而然就认为这对于计算机来说肯定也是小菜一碟;但是很快,人们意识到这种想法太乐观了。

如何才能在没有特定说明或规则的情况下识别数字呢?这似乎很难回

答。但是,人们更关心的是:为什么很难回答呢?因为我们从未学到过这样的规则。从小的时候,我们仅仅学过“这个是0”,“那个是1”。我们只是在思考这些数字是什么,并且在见过大量数字图片之后,能够识别得越来越好。是这样吧?那么,计算机又该如何做呢?不妨让计算机也重复相同的过程!是的,祝贺你,你现在抓住了机器学习的概念核心。机器学习被用来解决那些直接使用解析式很难解决的问题。机器学习建立模型的核心思想是在不容易建立公式和规则的情况下,使用训练的数据“通过合适的算法构建出一个模型”。

1.3 机器学习的挑战

已经知道,机器学习是一种从数据中找到(或者称为“学习”)模型参数的技术。因此,它很适合涉及“智慧”的问题,如图像识别和语音识别,在这些问题中,通过物理规则或数学公式几乎无法得到模型。机器学习使用的方法,一方面使得这个过程是可行的;但另一方面,又带来了更严重的问题。本节讨论机器学习面对的基础问题。

一旦机器学习根据训练数据得到了模型,人们就可以在实际数据中使用这个模型。这个过程可用图1-5来表示。图中垂直的箭头表示学习过程,即建立模型的过程,横向的箭头表示预测过程。

用于建立机器学习模型的训练数据与用于预测的输入数据是不同的。下面通过在图1-5中加入一项内容来更好地解释这个问题,如图1-6所示。

训练数据与输入数据的不同给机器学习带来了巨大的挑战。毫不夸张地说,机器学习的所有问题都来自于此。比如,使用某个人的所有手写体数