

深度学习 算法实践

(基于Theano和TensorFlow)

闫涛 周琦 编著



深度学习 算法实践

(基于Theano和TensorFlow)

闫涛 周琦 编著



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书以深度学习算法入门为主要内容,通过系统介绍 Python、NumPy、SciPy 等科学计算库,深度学习主流算法,深度学习前沿研究,深度学习服务云平台构建四大主线,向读者系统地介绍了深度学习的主要内容和研究进展。本书介绍了 Python、NumPy、SciPy 的使用技巧,面向谷歌推出的开源深度学习框架 TensorFlow,向读者展示了利用 TensorFlow 和 Theano 框架实现线性回归、逻辑回归、多层感知器、卷积神经网络、递归神经网络、长短时记忆网络、去噪自动编码器、堆叠自动编码器、受限玻尔兹曼机、深度信念网络等,并将这些技术用于 MNIST 手写数字识别任务。本书不仅讲述了深度学习算法本身,而且重点讲述了如何将这些深度学习算法包装成 Web 服务。本书旨在帮助广大工程技术人员快速掌握深度学习相关理论和实践,并将这些知识应用到实际工作中。

本书可以作为各类深度学习培训班的教材,也可以作为全国高等工科院校“深度学习”课程的教材,还可以作为广大人工智能、深度学习领域工程技术人员的参考书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

深度学习算法实践:基于 Theano 和 TensorFlow / 闫涛,周琦编著. —北京:电子工业出版社,2018.4
ISBN 978-7-121-33793-2

I. ①深… II. ①闫… ②周… III. ①人工智能—算法—研究 IV. ①TP18

中国版本图书馆 CIP 数据核字(2018)第 041449 号

策划编辑:付 睿

责任编辑:牛 勇

特约编辑:赵树刚

印 刷:北京京师印务有限公司

装 订:北京京师印务有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编:100036

开 本:787×1092 1/16 印张:36.5 字数:934.4 千字

版 次:2018 年 4 月第 1 版

印 次:2018 年 4 月第 1 次印刷

定 价:109.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888, 88258888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式:010-51260888-819, faq@phei.com.cn。

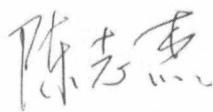
推荐序

《深度学习算法实践（基于 Theano 和 TensorFlow）》针对深度学习初学者的需求，详细讲解了深度学习典型算法的数学原理，给出了基于 TensorFlow 和 Theano 的算法实现，并以手写数字识别、图像标注、文本生成等为例，演示了深度学习算法的典型应用。作者立足于引导读者从解决问题的思路出发，层层剖析，逐步开发出实用的深度学习系统。通过阅读本书，在熟练掌握深度学习基本数学原理的基础上，读者不仅可以直接将书中内容用于项目实践，而且可以跟踪理解深度学习的最新进展。

自 2017 年下半年以来，深度学习又有了一些新进展，如注意力机制、生成式对抗网络、胶囊网络等，虽然本书还没有将其详细纳入，但是相信读者基于本书的知识架构，通过阅读相关论文及文献，理解并掌握这些算法并不困难。

当前，人工智能、深度学习技术的发展可谓一日千里，需要时刻跟踪业界的最新进展，才能保证自己的知识结构跟上业界发展步伐。作者拥有较深的理论造诣和丰富的实践经验，希望本书能够帮助读者掌握完整的知识体系，拥有较强的动手能力，成为人工智能、深度学习领域的学习型和实践型人才。

中国工程院院士



2018年3月9日

轻松注册成为博文视点社区用户（www.broadview.com.cn），扫码直达本书页面。

- **提交勘误：**您对书中内容的修改意见可在 [提交勘误](#) 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动：**在页面下方 [读者评论](#) 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/33793>



前 言

2016年上半年，随着 AlphaGo 战胜围棋世界冠军李世石，深度学习技术迅速进入大众的视野，成为热门技术。无论是科技领域还是创业投资领域，深度学习技术都受到了前所未有的关注，面向深度学习技术的初创公司不断刷新融资数额的纪录。

随着深度学习技术的流行，市场上对于深度学习人才的需求量激增。但是，由于深度学习技术涉及高等数学、线性代数、数理统计和机器学习相关知识，故学习曲线比较陡峭。目前国内外著名大学深度学习相关专业硕博士、国家重点实验室学生，都被 BAT 等一线互联网公司抢光了，人才大战有愈演愈烈之势。普通公司和初创公司，在这场人才大战中处于劣势，面临着越来越严重的人才荒。

随着深度学习技术的深入应用，企业对深度学习技术人才的渴求是正常的，但是当前市场上对深度学习技术人才的需求是非理性的。一方面，大家疯抢的深度学习人才都是著名院校的硕博士和重点实验室的研究员，但是这部分人所受的训练和精通的领域是做算法研究，而企业的目的是实际应用，二者在很大程度上是不匹配的，最后可能是企业下了血本却没有产生效益；另一方面，对于很多深度学习应用来说，对项目本身业务的理解，比对深度学习算法的理解要重要得多，而由于业务专家不懂深度学习技术，致使很多好的应用领域没有合适的人才来做。

编写本书的目的就是解决上述问题，希望读者可以通过对本书的学习，迅速掌握深度学习的理论框架和知识体系，具备在自己的专业领域内应用深度学习技术的能力，同时还具备跟踪深度学习领域最新进展的能力，能够独立复现顶级期刊文章中介绍的新方法和新理论。

本书内容及知识体系

第一部分为深度学习算法概述，包括第 1 章。

第 1 章简单介绍神经网络和深度学习发展史、现状和发展趋势，介绍并比较了深度学习开源框架，还介绍了开源框架的选择标准。

第二部分为深度学习算法基础，主要讲述深度学习算法中比较成熟的算法，包括第 2 章到第 7 章。

第 2 章介绍 Python 开发环境的搭建、NumPy 的使用、Theano 的安装和使用，并用本章介绍的知识实现一个简单的线性回归算法的程序。

第 3 章讲述逻辑回归算法的数学推导过程，并且讲述了通用学习模型的相关知识，还利用逻辑回归算法对 MNIST 手写数字识别数据集进行训练和识别。

第 4 章讲述多层感知器模型算法推导过程，以及数值计算与向量表示方式，并利用多层感知器模型对 MNIST 手写数字识别数据集进行训练和识别。

第 5 章讲述卷积神经网络的数学原理，详细讲解卷积神经网络的层间稀疏连接、权值共享和最大池化等技术，并利用卷积神经网络模型对 MNIST 手写数字识别数据集进行训练和识别。

第 6 章讲述递归神经网络的数学原理，以字符 RNN 网络为例，向读者演示了简单的计算机写作系统。而且，以微软图像标注数据集为例，以测试驱动开发的形式，向读者介绍利用递归神经网络做图像标注的典型应用。

第 7 章讲述长短时记忆网络的网络架构和数学原理，并以大型影评数据集为例，采用长短时记忆网络进行情感计算。

第三部分为深度学习算法进阶，主要讲述深度学习算法中比较前沿的算法，包括第 8 章到第 11 章。

第 8 章讲述自动编码器的数学原理，重点介绍实际中应用较多的去噪自动编码器和稀疏自动编码器，并以去噪自动编码器为例，对 MNIST 手写数字识别数据集进行特征提取。

第 9 章讲述将去噪自动编码器进行分层训练，组合成堆叠去噪自动编码器，并将堆叠去噪自动编码器用于 MNIST 手写数字识别任务。

第 10 章讲述受限玻尔兹曼机的数学原理，并将其用于 MNIST 手写数字识别任务。

第 11 章讲述深度信念网络的数学原理，以及其与受限玻尔兹曼机的关系，并将其用于 MNIST 手写数字识别任务。

第四部分为机器学习基础，主要讲述一些基础的机器学习算法，包括第 12 章和第 13 章。

第 12 章讲述生成式学习的基础理论，并将高斯判别分析用于癌症判别，将朴素贝叶斯算法用于垃圾邮件过滤。

第 13 章简单介绍支撑向量机算法的数学原理。

第五部分为深度学习平台 API，这部分讲述将深度学习算法包装成深度学习服务云平台的技术，包括第 14 章和第 15 章。

第 14 章介绍 Python 的 Web 开发环境及开发技术。

第 15 章应用 Web 开发技术，将前面介绍的多层感知器模型包装成 RESTful 服务，用户可以通过网页上传图片文件，并得到识别后的结果。

由于篇幅所限，书中很多例子只给出了部分代码，这些代码对于理解算法的实现原理是足够的，但是考虑到代码的完整性，我们将书中绝大部分例程都上传到了 GitHub 的开源项目 <https://github.com/yt7589/dlp.git>，书中的代码放在 book/chp** 目录下，这些代码在 Ubuntu 16.04+Python3.6+TensorFlow1.2 和 Windows+Anaconda+TensorFlow1.2 下均可正常运行。读者可以下载相关源码，通过运行这些源码加深对书中内容的理解。

目 录

第一部分 深度学习算法概述

第 1 章 深度学习算法简介.....	2
1.1 神经网络发展简史.....	2
1.1.1 神经网络第一次兴起.....	3
1.1.2 神经网络沉寂期（20 世纪 80 年代—21 世纪）.....	4
1.1.3 神经网络技术积累期（20 世纪 90 年代—2006 年）.....	5
1.1.4 深度学习算法崛起（2006 年至今）.....	8
1.2 深度学习现状.....	10
1.2.1 传统神经网络困境.....	10
1.2.2 深度多层感知器.....	12
1.2.3 深度卷积神经网络.....	14
1.2.4 深度递归神经网络.....	15
1.3 深度学习研究前瞻.....	16
1.3.1 自动编码器.....	17
1.3.2 深度信念网络.....	18
1.3.3 生成式网络最新进展.....	19
1.4 深度学习框架比较.....	20
1.4.1 TensorFlow.....	20
1.4.2 Theano.....	21
1.4.3 Torch.....	22
1.4.4 DeepLearning4J.....	23
1.4.5 Caffe.....	23
1.4.6 MXNet.....	24
1.4.7 CNTK.....	27
1.4.8 深度学习框架造型指导原则.....	27
1.5 深度学习入门路径.....	28
1.5.1 运行 MNIST.....	28

1.5.2	深度学习框架的选择.....	29
1.5.3	小型试验网络.....	33
1.5.4	训练生产网络.....	33
1.5.5	搭建生产环境.....	34
1.5.6	持续改进.....	35

第二部分 深度学习算法基础

第 2 章	搭建深度学习开发环境.....	38
2.1	安装 Python 开发环境.....	38
2.1.1	安装最新版本 Python.....	38
2.1.2	Python 虚拟环境配置.....	39
2.1.3	安装科学计算库.....	40
2.1.4	安装最新版本 Theano.....	40
2.1.5	图形绘制.....	40
2.2	NumPy 简易教程.....	43
2.2.1	Python 基础.....	43
2.2.2	多维数组的使用.....	51
2.2.3	向量运算.....	58
2.2.4	矩阵运算.....	60
2.2.5	线性代数.....	62
2.3	TensorFlow 简易教程.....	68
2.3.1	张量定义.....	69
2.3.2	变量和 placeholder.....	69
2.3.3	神经元激活函数.....	71
2.3.4	线性代数运算.....	72
2.3.5	操作数据集.....	74
2.4	Theano 简易教程.....	77
2.4.1	安装 Theano.....	77
2.4.2	Theano 入门.....	78
2.4.3	Theano 矩阵相加.....	79
2.4.4	变量和共享变量.....	80
2.4.5	随机数的使用.....	84
2.4.6	Theano 求导.....	84
2.5	线性回归.....	86
2.5.1	问题描述.....	86
2.5.2	线性模型.....	88

2.5.3	线性回归学习算法.....	89
2.5.4	解析法.....	90
2.5.5	Theano 实现.....	93
第 3 章	逻辑回归	100
3.1	逻辑回归数学基础.....	100
3.1.1	逻辑回归算法的直观解释.....	100
3.1.2	逻辑回归算法数学推导.....	101
3.1.3	牛顿法解逻辑回归问题.....	103
3.1.4	通用学习模型.....	106
3.2	逻辑回归算法简单应用.....	113
3.3	MNIST 手写数字识别库简介	124
3.4	逻辑回归 MNIST 手写数字识别	126
第 4 章	感知器模型和 MLP	139
4.1	感知器模型.....	139
4.1.1	神经元模型.....	139
4.1.2	神经网络架构.....	143
4.2	数值计算形式.....	144
4.2.1	前向传播.....	144
4.2.2	误差反向传播.....	145
4.2.3	算法推导.....	147
4.3	向量化表示形式.....	152
4.4	应用要点.....	153
4.4.1	输入信号模型.....	154
4.4.2	权值初始化.....	155
4.4.3	早期停止.....	155
4.4.4	输入信号调整.....	156
4.5	TensorFlow 实现 MLP	156
第 5 章	卷积神经网络	174
5.1	卷积神经网络原理.....	174
5.1.1	卷积神经网络的直观理解.....	174
5.1.2	卷积神经网络构成.....	177
5.1.3	卷积神经网络设计.....	191
5.1.4	迁移学习和网络微调.....	193
5.2	卷积神经网络的 TensorFlow 实现.....	195
5.2.1	模型搭建.....	197

5.2.2	训练方法.....	203
5.2.3	运行方法.....	208
第 6 章	递归神经网络.....	212
6.1	递归神经网络原理.....	212
6.1.1	递归神经网络表示方法.....	213
6.1.2	数学原理.....	214
6.1.3	简单递归神经网络应用示例.....	219
6.2	图像标记.....	226
6.2.1	建立开发环境.....	226
6.2.2	图像标记数据集处理.....	227
6.2.3	单步前向传播.....	229
6.2.4	单步反向传播.....	231
6.2.5	完整前向传播.....	234
6.2.6	完整反向传播.....	236
6.2.7	单词嵌入前向传播.....	239
6.2.8	单词嵌入反向传播.....	241
6.2.9	输出层前向/反向传播.....	243
6.2.10	输出层代价函数计算.....	245
6.2.11	图像标注网络整体架构.....	248
6.2.12	代价函数计算.....	249
6.2.13	生成图像标记.....	255
6.2.14	网络训练过程.....	258
6.2.15	网络持久化.....	265
第 7 章	长短时记忆网络.....	269
7.1	长短时记忆网络原理.....	269
7.1.1	网络架构.....	269
7.1.2	数学公式.....	272
7.2	MNIST 手写数字识别.....	274

第三部分 深度学习算法进阶

第 8 章	自动编码器.....	286
8.1	自动编码器概述.....	286
8.1.1	自动编码器原理.....	287
8.1.2	去噪自动编码器.....	287

8.1.3 稀疏自动编码器.....	288
8.2 去噪自动编码器 TensorFlow 实现.....	291
8.3 去噪自动编码器的 Theano 实现.....	298
第 9 章 堆叠自动编码器	307
9.1 堆叠去噪自动编码器.....	308
9.2 TensorFlow 实现.....	322
9.3 Theano 实现.....	341
第 10 章 受限玻尔兹曼机	344
10.1 受限玻尔兹曼机原理.....	344
10.1.1 网络架构.....	344
10.1.2 能量模型.....	346
10.1.3 CD-K 算法.....	351
10.2 受限玻尔兹曼机 TensorFlow 实现.....	353
10.3 受限玻尔兹曼机 Theano 实现.....	362
第 11 章 深度信念网络	381
11.1 深度信念网络原理.....	381
11.2 深度信念网络 TensorFlow 实现.....	382
11.3 深度信念网络 Theano 实现.....	403

第四部分 机器学习基础

第 12 章 生成式学习.....	420
12.1 高斯判别分析.....	422
12.1.1 多变量高斯分布.....	422
12.1.2 高斯判决分析公式.....	423
12.2 朴素贝叶斯.....	436
12.2.1 朴素贝叶斯分类器.....	436
12.2.2 拉普拉斯平滑.....	439
12.2.3 多项式事件模型.....	441
第 13 章 支撑向量机.....	444
13.1 支撑向量机概述.....	444
13.1.1 函数间隔和几何间隔.....	445
13.1.2 最优距离分类器.....	448
13.2 拉格朗日对偶.....	448

13.3	最优分类器算法.....	450
13.4	核方法.....	453
13.5	非线性可分问题.....	455
13.6	SMO 算法.....	457
13.6.1	坐标上升算法.....	458
13.6.2	SMO 算法详解.....	458

第五部分 深度学习平台 API

第 14 章	Python Web 编程.....	462
14.1	Python Web 开发环境搭建.....	462
14.1.1	CherryPy 框架.....	463
14.1.2	CherryPy 安装.....	463
14.1.3	测试 CherryPy 安装是否成功.....	464
14.2	最简 Web 服务器.....	465
14.2.1	程序启动.....	465
14.2.2	显示 HTML 文件.....	466
14.2.3	静态内容处理.....	468
14.3	用户认证系统.....	471
14.4	AJAX 请求详解.....	473
14.4.1	添加数据.....	474
14.4.2	修改数据.....	476
14.4.3	删除数据.....	478
14.4.4	REST 服务实现.....	479
14.5	数据持久化技术.....	487
14.5.1	环境搭建.....	487
14.5.2	数据库添加操作.....	488
14.5.3	数据库修改操作.....	489
14.5.4	数据库删除操作.....	490
14.5.5	数据库查询操作.....	491
14.5.6	数据库事务操作.....	492
14.5.7	数据库连接池.....	494
14.6	任务队列.....	499
14.7	媒体文件上传.....	502
14.8	Redis 操作.....	504
14.8.1	Redis 安装配置.....	504
14.8.2	Redis 使用例程.....	505

第 15 章 深度学习云平台	506
15.1 神经网络持久化.....	506
15.1.1 数据库表设计.....	506
15.1.2 整体目录结构.....	511
15.1.3 训练过程及模型文件保存.....	512
15.2 神经网络运行模式.....	528
15.3 AJAX 请求调用神经网络	531
15.3.1 显示静态网页.....	531
15.3.2 上传图片文件.....	540
15.3.3 AJAX 接口	543
15.4 请求合法性验证.....	545
15.4.1 用户注册和登录.....	546
15.4.2 客户端生成请求.....	553
15.4.3 服务器端验证请求.....	555
15.5 异步结果处理.....	557
15.5.1 网页异步提交.....	557
15.5.2 应用队列管理模块.....	559
15.5.3 任务队列.....	560
15.5.4 结果队列.....	561
15.5.5 异步请求处理流程.....	562
15.6 神经网络持续改进.....	563
15.6.1 应用遗传算法.....	563
15.6.2 重新训练.....	564
15.6.3 生成式对抗网络.....	565
后 记	567
参考文献.....	568

章 第

介绍深度学习算法

第一部分 深度学习算法概述

深度学习算法概述

□ 深度学习算法简介

第 1 章

深度学习算法简介

1.1 神经网络发展简史

人工神经网络 (ANN) 是由许多简单的、相互连接的处理单元组成的, 这些处理单元被称为神经元。人工神经网络是对生物神经系统进行仿生设计的结果, 因此人工神经元与生物神经元有许多相似之处, 每个神经元都具有若干个输入信号, 神经元经过处理 (通常是非线性处理) 产生一个实数值的输出, 通过输出突触传递给周围的神经元。整个人工神经网络, 通过接收外界的输入信号, 经过神经元的协同运算, 产生人们希望的结果。从过去几十年的发展历史来看, 相对于其他方法而言, 人工神经网络对于很难用显性知识表示的领域, 例如计算机视觉、模式识别、人脸识别、语音识别、视频识别等领域, 具有非常大的优势。

神经网络的学习过程主要是指通过发现合适的神经元间的连接权值, 来使整个神经网络表现出我们希望的行为, 如自动驾驶、图像识别等。根据需要处理的问题, 以及神经元间连接方式的不同, 神经网络的决策过程可能包括很长的因果链, 或者说多个计算阶段。在每个阶段中, 神经网络都会自动汇总网络的激活信号, 经过特定的非线性变换, 为下一阶段的运算做好准备, 并最终产生我们需要的结果。深度学习网络正是这样一种结构, 它完美地体现了分阶段任务的特性, 因此在实际应用中取得了令人刮目相看的成就。

我们知道, 浅层神经网络在很久以前就存在了, 20 世纪 60、70 年代, 就有多层非线性神经网络应用的例子。而且, 基于监督学习的梯度下降方法, 对于连续可微的多重函数, 人们发现了反向传播算法 (BP), 并且在 1981 年成功将其应用于神经网络, 使得基于这种算法的前馈神经网络在 20 世纪 80 年代迅速流行起来。然而在 20 世纪 80 年代末期, 人们发现将 BP 算法应用于深度网络会出现非常大的问题, 鉴于即使只有一个隐藏层神经网络,

只要隐藏层神经元数目足够多，也可以拟合任意函数，因此当时人们普遍认为浅而宽的神经网络具有更好的性能。但是对于深度网络的研究并没有停止，尤其是研究人员将非监督学习引入深度学习网络之后，随着算法和神经元激活函数方面的一系列改进，使得深度学习网络在一系列模式识别竞赛中获奖，尤其是2011年，在某些特定领域，深度学习网络的模式识别能力甚至超过了人类。与此同时，深度学习在商业应用方面也取得了巨大的成功，例如在图像识别、语音识别和自然语言处理方面，取得了一系列令人瞩目的成绩，这一趋势一直延续至今。

在过去十几年中，前向神经网络（FNN）和递归神经网络（RNN）都赢得了一系列的算法竞赛。从某种意义上来说，递归神经网络是最深的神经网络，拥有比前向神经网络更强的计算能力，因为递归神经网络将序列信号分析和信息并行处理，以一种自然、高效的方式结合在一起。配合当前不断增加的计算能力，可以预见递归神经网络将拥有越来越重要的地位。

然而，神经网络的发展并不是一帆风顺的，中间经历数次大起大落，本节我们将简要回顾一下神经网络的发展史，通过介绍深度学习之前神经网络所遇到的困境，使读者对深度学习的背景有一个清晰的了解，这对于理解深度学习发展趋势是非常重要的，因为历史总是惊人的相似。同时，我们也应该看到，深度学习并不是人工神经网络发展的终点，其只是人工神经网络发展的一个特定阶段，一定会有更先进的技术超越当前的深度学习技术，因为深度学习网络目前只是将学习过程进行了分解，同时用非监督学习来辅助监督学习，并没有从根本上解决人工神经网络中欠拟合（Under Fitting）和过拟合（Over Fitting）等泛化问题。如果可以在强化学习和非监督学习方面取得突破，理论上可以产生更好的学习算法。

1.1.1 神经网络第一次兴起

现代神经网络起源于20世纪40年代，早期典型的神经网络模型有Hebb在1949年提出的非监督学习模型，以及1958年Rosenblatt提出的基于监督学习的感知器模型，其后出现了一系列监督学习和非监督学习模型。总体来讲，由于这一阶段计算能力有限，神经网络模型都比较简单，而且缺乏成功的实际应用。

最早的深度学习网络模型出现在20世纪60~70年代，Ivakhnenko等人提出了GMDH神经网络模型，在这种网络模型下，神经元的激活函数为Kolmogorov-Gabor多项式。在给定一个训练集的情况下，先通过回归分析增加神经网络的层数，再通过单独的验证集对网络拓扑结构进行剪裁。综上所述，GMDH神经网络的网络层数和每层的神经元数，都可以通过学习来进行调整，并且针对特定的问题，网络拓扑结构也将不同。可以说，GMDH神经网络是第一种开放的、分层学习的神经网络模型，当时最深的网络可以达到8层。

在20世纪60~70年代，神经生理学家们发现，在猫的视觉皮层中有两类细胞，一类是简单细胞，另一类是复杂细胞。简单细胞对图像中的细节信息更敏感，例如边缘等；