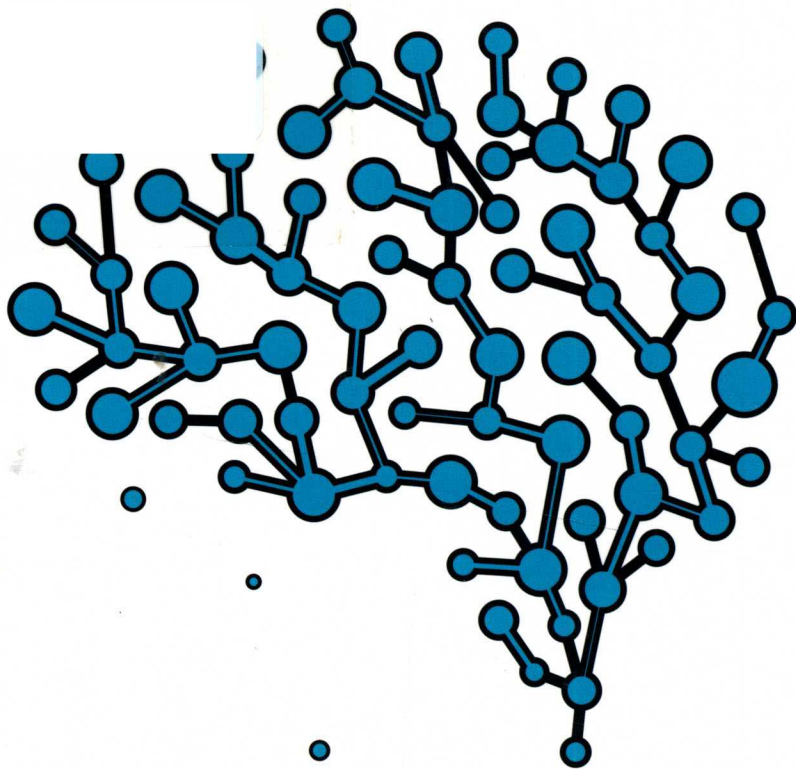


TURING

图灵原创



# 卷积神经网络的 Python实现

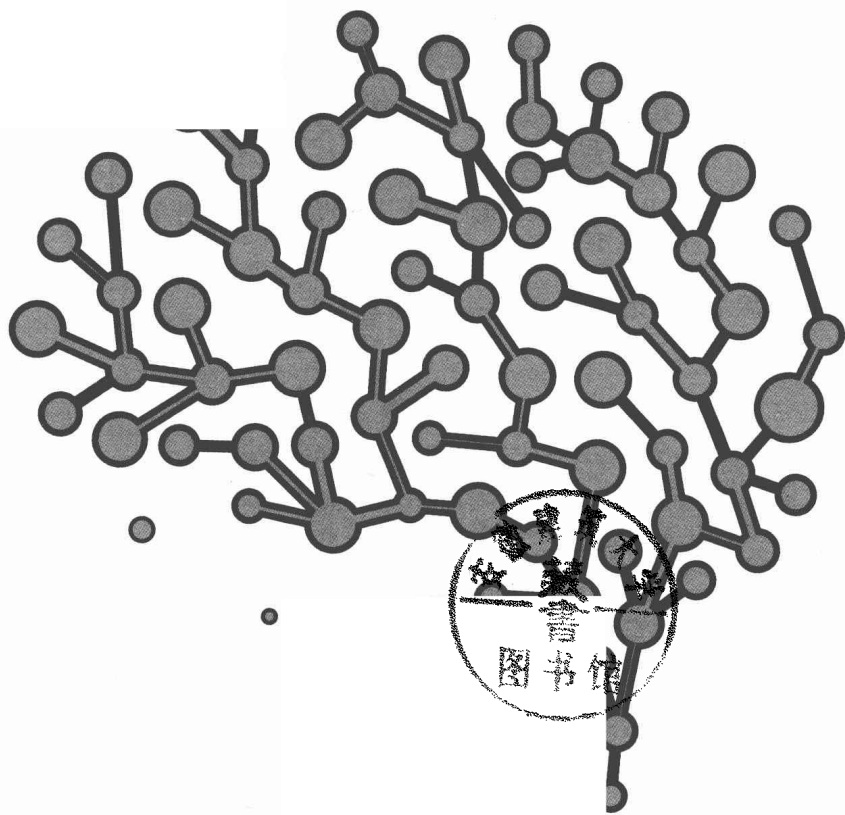
单建华◎著



中国工信出版集团



人民邮电出版社  
POSTS & TELECOM PRESS



# 卷积神经网络的 Python实现

单建华◎ 著

人民邮电出版社  
北京

## 图书在版编目 (C I P) 数据

卷积神经网络的Python实现 / 单建华著. -- 北京 :  
人民邮电出版社, 2019. 1  
(图灵原创)  
ISBN 978-7-115-49756-7

I. ①卷… II. ①单… III. ①人工神经网络—软件工  
具—程序设计 IV. ①TP183

中国版本图书馆CIP数据核字(2018)第239530号

## 内 容 提 要

卷积神经网络是深度学习最重要的模型之一。本书是卷积神经网络领域的入门读物, 假定读者不具备任何机器学习知识。书中尽可能少地使用数学知识, 从机器学习的概念讲起, 以卷积神经网络的最新发展结束。

本书首先简单介绍了机器学习的基本概念, 详细讲解了线性模型、神经网络和卷积神经网络模型, 然后介绍了基于梯度下降法的优化方法和梯度反向传播算法, 接着介绍了训练网络前的准备工作、神经网络实战、卷积神经网络的应用及其发展。针对每个关键知识点, 书中给出了基于NumPy的代码实现, 以及完整的神经网络和卷积神经网络代码实现, 方便读者训练网络和查阅代码。

本书既可以作为卷积神经网络的教材, 也可以供对卷积神经网络感兴趣的工程技术人员和科研人员参考。

- 
- ◆ 著 单建华  
责任编辑 王军花  
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>  
三河市君旺印务有限公司印刷
  - ◆ 开本: 880×1230 1/32  
印张: 7.5 彩插: 2  
字数: 166千字 2019年1月第1版  
印数: 1-3 000册 2019年1月河北第1次印刷
- 

定价: 49.00元

读者服务热线: (010)51095186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

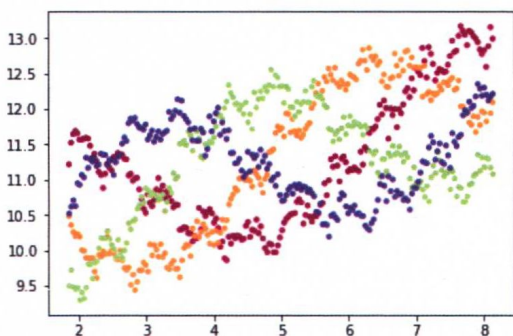


图 8.1 该数据集有 4 类样本，是线性增长的振荡数据

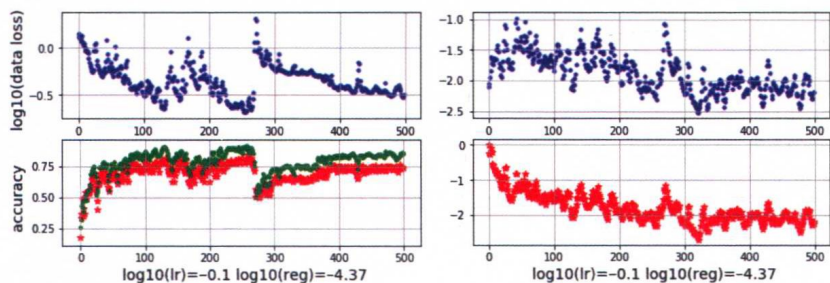


图 8.2 学习率过大，数据损失波动很大且下降很快，参数更新率也过高

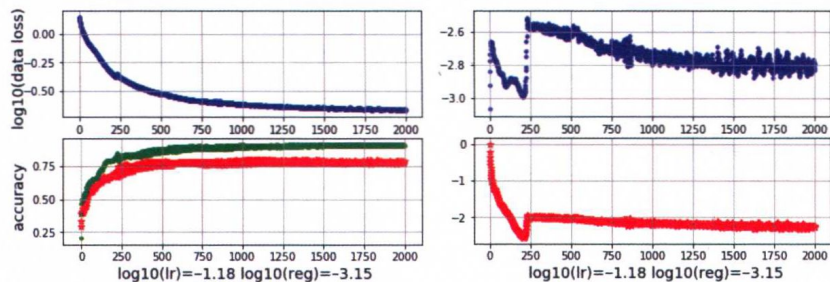


图 8.3 数据损失无波动，训练集和验证集的准确率饱和，验证集的准确率只有 75%，参数更新率稍大，这说明减小学习率有可能进一步提高验证集的准确率

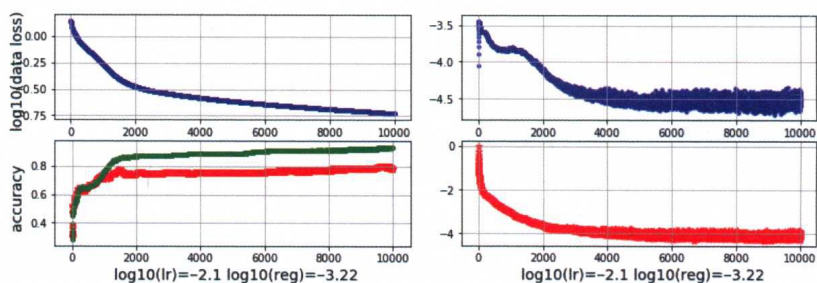


图 8.4 数据损失无波动，训练集和验证集的准确率趋于饱和，验证集的准确率有 80%，训练集的准确率高达 95%，收敛时参数更新率已经很低，过拟合现象不明显，这说明模型验证集的准确率难以超过 80%

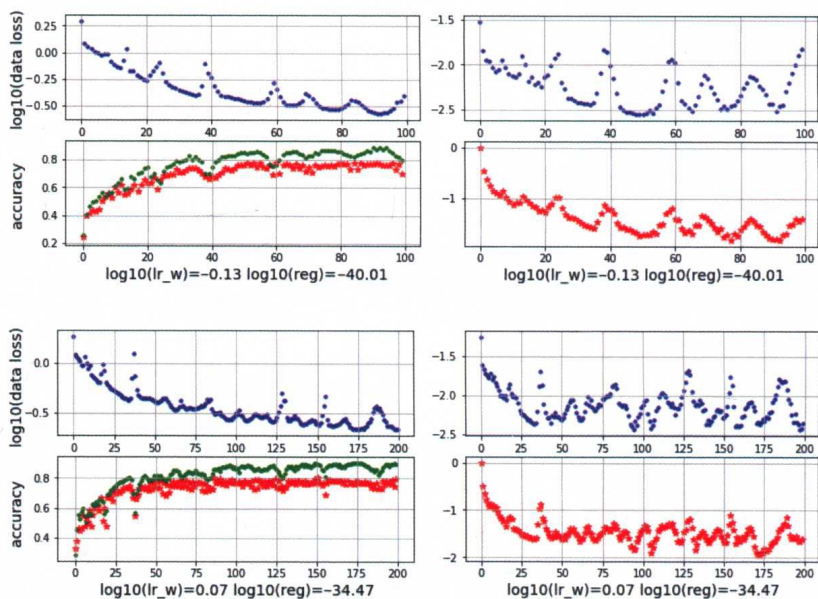


图 8.5 很大的初始学习率和很小的正则化，达到很好的性能

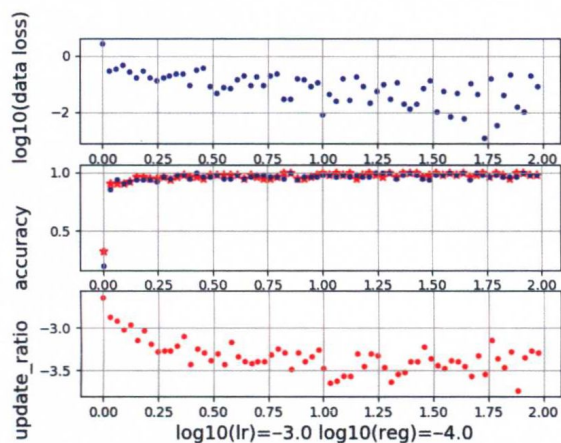


图 9.1 训练过程

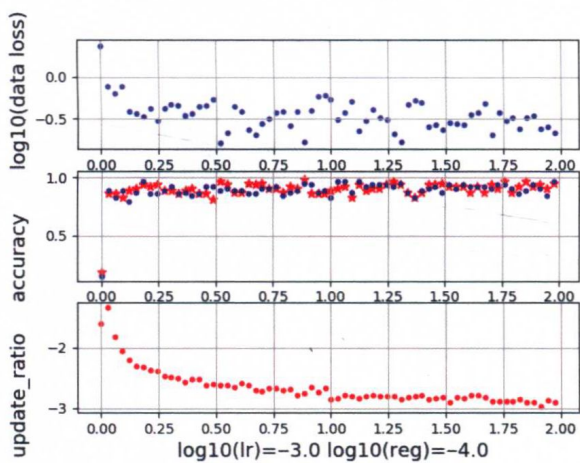


图 9.2 线性模型的训练过程

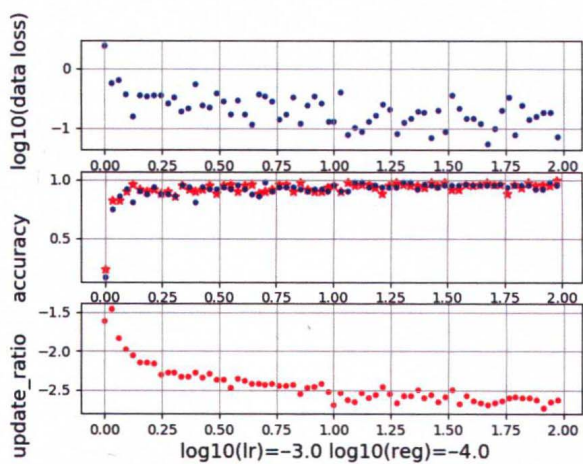


图 9.3 一个隐含层（64 个神经元）神经网络模型的训练过程

# 前 言

近几年，深度学习在学术界和工业界掀起革命，Python 语言作为机器学习的首选语言也异军突起。然而，大多数 Python 语言的教材侧重语法，很少有项目实战，深度学习方面的教材侧重理论，特别是针对卷积神经网络，很少涉及源码解读和实现。因此我试图通过本书，使读者掌握 Python 语言并深入理解卷积神经网络，并能用 Python 语言实现卷积神经网络结构。

为了使尽可能多的读者对卷积神经网络有所了解，本书假定读者不具备任何机器学习知识。同时，我尽量少地使用数学知识，只要掌握向量、矩阵及矩阵乘法的定义，就能理解本书中除了优化之外的所有章节；优化问题涉及的数学知识有导数、偏导数、泰勒展开式和链式法则。学习 Python 语言需要掌握一些基本知识，如 list 结构、for 循环、函数、类和 NumPy 模块。因此，本书适合具有一定 Python 基础的大学二年级以上的理工科本科生和研究生，具有类似知识基础的工程技术人员和科研人员，以及对卷积神经网络有一定了解，想进一步理解的人士阅读。



全书共 10 章，分为 3 个部分：第一部分（第 1 章~第 4 章）介绍了机器学习的基本概念、线性模型、神经网络和卷积神经网络模型，采用一条主线贯穿这 3 种模型，层层递进，降低了学习难度；第二部分（第 5 章和第 6 章）介绍了基于梯度下降法的优化方法和梯度反向传播算法，详细地给出了各种模块的反向传播代码，读者掌握了这些方法后，如果遇到新的模块，就能独立实现其反向传播代码；第三部分（第 7 章~第 10 章）介绍了训练网络前的准备工作、神经网络实战、卷积神经网络的应用及其发展，特别详细地给出了数据归一化、梯度检测、批量归一化(BN)及各种经典的卷积神经网络结构，如 VGGNet、GoogLeNet、ResNet、SENet 和轻量网络 MobileNetV2。除第 1 章外，每个关键知识点都给出了基于 NumPy 的代码实现，特别是第 8 章和第 9 章这两章，给出了完整的神经网络和卷积神经网络代码实现，方便读者训练网络和查阅代码。

如果你只是想了解卷积神经网络的原理，可只看第一部分和第 10 章。如果你只是想更好地利用 TensorFlow 等框架来训练网络结构，可不用看第 6 章，通过学习第 5 章就能正确选择合适的优化算法。如果你想不借助 TensorFlow 等框架实现各种网络结构，则需学习所有章节，特别是第二部分。

本书中的代码<sup>①</sup>的一个突出优点是接近伪代码，几乎不言自明，非常容易读懂。如果读者没有 Python 基础，不看代码也能掌握卷积神经网络的知识。当然，强烈建议读者手工输入代码，泛读或精读代码，

---

① 本书代码可以从图灵社区(iTuring.cn)本书主页免费注册下载。

注意其实现细节。这能加深对卷积神经网络的理解，同时也是一个极好的 Python 语言项目实战训练。

卷积神经网络发展极为迅速，各种新网络结构层出不穷，理论也在一步一步发展，本书只是对卷积神经网络最基本、最核心的概念，以及最重要的网络结构进行介绍。我自认才疏学浅，略知皮毛，更兼精力有限，书中错谬之处在所难免，若蒙读者不吝告知，将不胜感激。

我的邮箱是 [shanjianhua.vip@qq.com](mailto:shanjianhua.vip@qq.com)，如有疑问，欢迎垂询。

最后要感谢我的家人，特别感谢我的爱人余慧莉女士，感谢她一直以来对我的理解和支持。当然，也要感谢使本书顺利出版的编辑王军花和武芮欣两位女士。

单建华

2018年5月于马鞍山

# 目 录

## 第一部分 模型篇

第 1 章 机器学习简介 .....	2
1.1 引言 .....	2
1.2 基本术语 .....	3
1.3 重要概念 .....	5
1.4 图像分类 .....	12
1.5 MNIST 数据集简介 .....	15
第 2 章 线性分类器 .....	17
2.1 线性模型 .....	17
2.1.1 线性分类器 .....	18
2.1.2 理解线性分类器 .....	19
2.1.3 代码实现 .....	21
2.2 softmax 损失函数 .....	22
2.2.1 损失函数的定义 .....	23
2.2.2 概率解释 .....	24
2.2.3 代码实现 .....	25
2.3 优化 .....	26

2.4	梯度下降法	26
2.4.1	梯度的解析意义	27
2.4.2	梯度的几何意义	29
2.4.3	梯度的物理意义	29
2.4.4	梯度下降法代码实现	29
2.5	牛顿法	30
2.6	机器学习模型统一结构	31
2.7	正则化	33
2.7.1	范数正则化	34
2.7.2	提前终止训练	37
2.7.3	概率的进一步解释	38
<b>第 3 章</b>	<b>神经网络</b>	<b>39</b>
3.1	数学模型	39
3.2	激活函数	41
3.3	代码实现	44
3.4	学习容量和正则化	45
3.5	生物神经科学基础	48
<b>第 4 章</b>	<b>卷积神经网络的结构</b>	<b>50</b>
4.1	概述	50
4.1.1	局部连接	51
4.1.2	参数共享	52
4.1.3	3D 特征图	52
4.2	卷积层	53
4.2.1	卷积运算及代码实现	54
4.2.2	卷积层及代码初级实现	57
4.2.3	卷积层参数总结	63
4.2.4	用连接的观点看卷积层	64
4.2.5	使用矩阵乘法实现卷积层运算	67
4.2.6	批量数据的卷积层矩阵乘法的代码实现	69

4.3	池化层	74
4.3.1	概述	74
4.3.2	池化层代码实现	76
4.4	全连接层	79
4.4.1	全连接层转化成卷积层	80
4.4.2	全连接层代码实现	82
4.5	卷积网络的结构	83
4.5.1	层的组合模式	83
4.5.2	表示学习	86
4.6	卷积网络的神经科学基础	87

## 第二部分 优化篇

第 5 章	基于梯度下降法的最优化方法	90
5.1	随机梯度下降法 SGD	91
5.2	基本动量法	93
5.3	Nesterov 动量法	95
5.4	AdaGrad	95
5.5	RMSProp	97
5.6	Adam	98
5.7	AmsGrad	99
5.8	学习率退火	99
5.9	参数初始化	100
5.10	超参数调优	101
第 6 章	梯度反向传播算法	104
6.1	基本函数的梯度	104
6.2	链式法则	105
6.3	深度网络的误差反向传播算法	107
6.4	矩阵化	109
6.5	softmax 损失函数梯度计算	111

6.6	全连接层梯度反向传播 .....	112
6.7	激活层梯度反向传播 .....	113
6.8	卷积层梯度反向传播 .....	115
6.9	最大值池化层梯度反向传播 .....	118

## 第三部分 实战篇

第 7 章	训练前的准备 .....	124
7.1	中心化和规范化 .....	124
7.1.1	利用线性模型推导中心化 .....	125
7.1.2	利用属性同等重要性推导规范化 .....	126
7.1.3	中心化和规范化的几何意义 .....	128
7.2	PCA 和白化 .....	128
7.2.1	从去除线性相关性推导 PCA .....	129
7.2.2	PCA 代码 .....	130
7.2.3	PCA 降维 .....	131
7.2.4	PCA 的几何意义 .....	133
7.2.5	白化 .....	134
7.3	卷积网络在进行图像分类时如何预处理 .....	135
7.4	BN .....	136
7.4.1	BN 前向计算 .....	136
7.4.2	BN 层的位置 .....	137
7.4.3	BN 层的理论解释 .....	138
7.4.4	BN 层在实践中的注意事项 .....	139
7.4.5	BN 层的梯度反向传播 .....	140
7.4.6	BN 层的地位探讨 .....	141
7.4.7	将 BN 层应用于卷积网络 .....	141
7.5	数据扩增 .....	142
7.6	梯度检查 .....	144
7.7	初始损失值检查 .....	146
7.8	过拟合微小数据集 .....	146

7.9 监测学习过程	147
7.9.1 损失值	147
7.9.2 训练集和验证集的准确率	148
7.9.3 参数更新比例	149
<b>第 8 章 神经网络实例</b>	<b>150</b>
8.1 生成数据	150
8.2 数据预处理	152
8.3 网络模型	153
8.4 梯度检查	156
8.5 参数优化	158
8.6 训练网络	159
8.7 过拟合小数据集	162
8.8 超参数随机搜索	162
8.9 评估模型	165
8.10 程序组织结构	165
8.11 增加 BN 层	167
8.12 程序使用建议	171
<b>第 9 章 卷积神经网络实例</b>	<b>172</b>
9.1 程序结构设计	173
9.2 激活函数	173
9.3 正则化	174
9.4 优化方法	175
9.5 卷积网络的基本模块	176
9.6 训练方法	181
9.7 VGG 网络结构	186
9.8 MNIST 数据集	197
9.9 梯度检测	199
9.10 MNIST 数据集的训练结果	202
9.11 程序使用建议	205

第 10 章 卷积网络结构的发展 .....	206
10.1 全局平均池化层 .....	206
10.2 去掉池化层 .....	208
10.3 网络向更深更宽发展面临的困难 .....	209
10.4 ResNet 向更深发展的代表网络 .....	210
10.5 GoogLeNet 向更宽发展的代表网络 .....	213
10.6 轻量网络 .....	215
10.6.1 $1 \times 1$ 深度维度卷积代码实现 .....	217
10.6.2 $3 \times 3$ 逐特征图的卷积代码实现 .....	219
10.6.3 逆残差模块的代码实现 .....	222
10.7 注意机制网络 SENet .....	223



## 第一部分

# 模型篇

- 第1章 机器学习简介
- 第2章 线性分类器
- 第3章 神经网络
- 第4章 卷积神经网络的结构