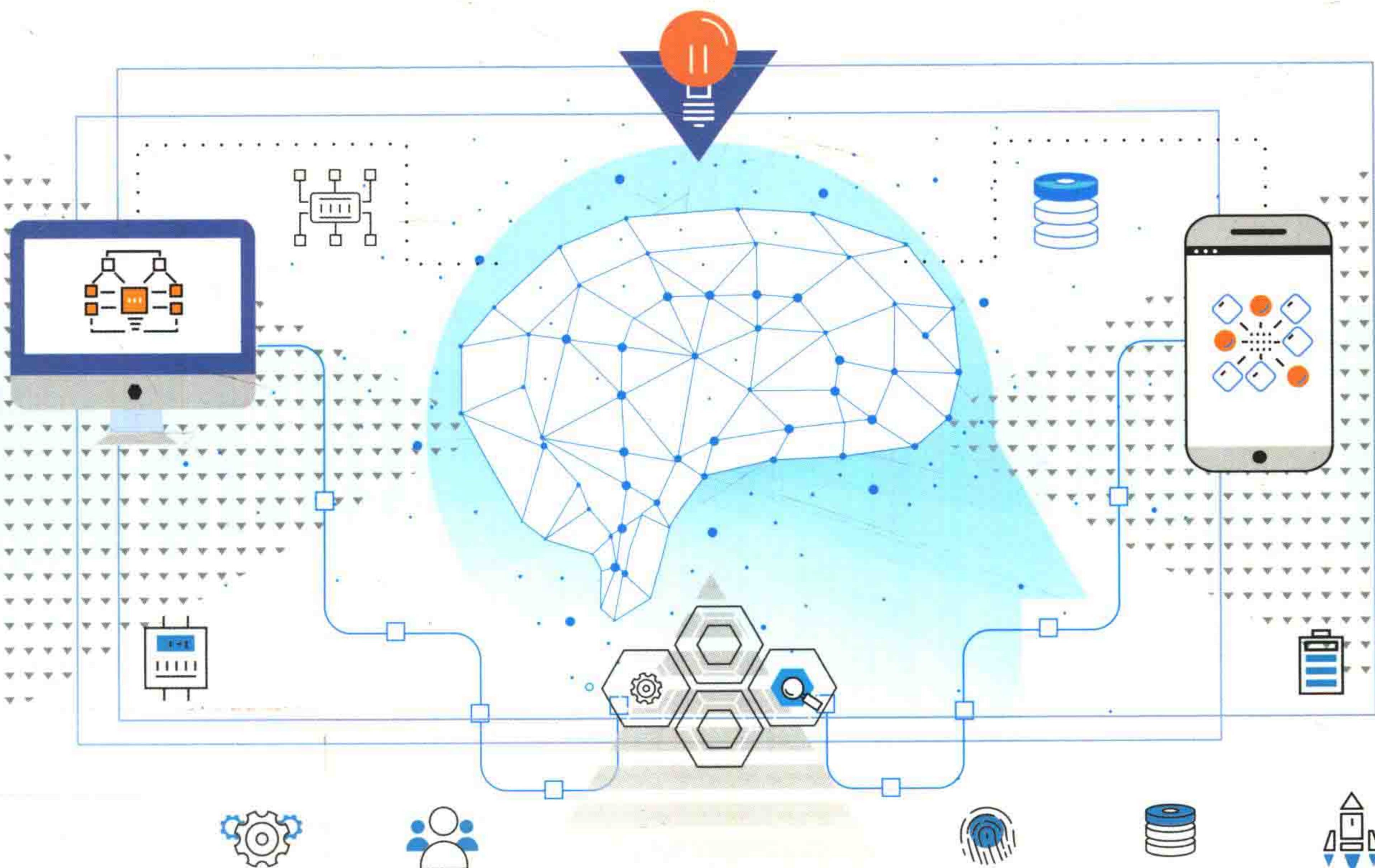


Neural Networks with R

神经网络 R语言实现

[美] 朱塞佩·查博罗 (Giuseppe Ciaburro) 著
巴拉伊·温卡特斯瓦兰 (Balaji Venkateswaran)
李洪成 译



机械工业出版社
China Machine Press

Neural Networks with R

神经网络

R语言实现

[美] 朱塞佩·查博罗 (Giuseppe Ciaburro) 著
巴拉伊·温卡特斯瓦兰 (Balaji Venkateswaran)
李洪成 译



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

神经网络: R 语言实现 / (美) 朱塞佩·查博罗 (Giuseppe Ciaburro), (美) 巴拉伊·温卡特斯瓦兰 (Balaji Venkateswaran) 著; 李洪成译. —北京: 机械工业出版社, 2018.7
(智能系统与技术丛书)

书名原文: Neural Networks with R

ISBN 978-7-111-60384-9

I. 神… II. ①朱… ②巴… ③李… III. 程序语言 - 程序设计 - 应用 - 人工神经网络
IV. TP183

中国版本图书馆 CIP 数据核字 (2018) 第 146530 号

本书版权登记号: 图字 01-2017-7512

Giuseppe Ciaburro, Balaji Venkateswaran: Neural Networks with R (ISBN: 978-1-78839-787-2).

Copyright © 2017 Packt Publishing. First published in the English language under the title “Neural Networks with R”.

All rights reserved.

Chinese simplified language edition published by China Machine Press.

Copyright © 2018 by China Machine Press.

本书中文简体字版由 Packt Publishing 授权机械工业出版社独家出版。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。

神经网络: R 语言实现

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 张梦玲

责任校对: 李秋荣

印 刷: 北京市荣盛彩色印刷有限公司

版 次: 2018 年 7 月第 1 版第 1 次印刷

开 本: 186mm × 240mm 1/16

印 张: 14.5

书 号: ISBN 978-7-111-60384-9

定 价: 59.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88379426 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzit@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

THE TRANSLATOR'S WORDS

译者序

神经网络的研究起源于 19 世纪后期和 20 世纪早期，它横跨工程学、物理学、神经生理学、心理学、计算机和数学等多个学科。早期的神经网络学者 Hermann Von Helmholtz 等强调学习理论、视觉等，并没有给出神经运算的特殊数学模型。

Warren McCulloch 在 20 世纪 40 年代给出了现代视角的神经网络，并有了神经网络的第一个实际应用。在 50 年代，开始有了感知机学习规则。但是人们之后发现，它仅能用于有限的应用问题。60 年代末，一些学者开始认为神经网络的进一步研究是没有前途的。80 年代初，人们又重新对神经网络产生了兴趣。进入 21 世纪后，由于计算机处理能力的增强，特别是 GPU 的计算能力增强，神经网络在图像处理和计算机视觉等领域取得了极大的成功。谷歌（Google）公司的 AlphaGo 应用深度神经网络的学习技术，战胜了当前世界上最优秀的围棋选手。近两年，学术界和工业界掀起了研究深度学习、深度神经网络的极大热潮。

神经网络跨越多个领域，不同领域的作者有时候采用不同的术语。因此，神经网络的基本概念虽然不是太复杂，但是相关的书籍阅读起来比较吃力，读者较难把握其中的概念。如果能够结合具体的实例并能够进行实际的操作，将会大大帮助读者掌握神经网络的基本知识和应用技能。本书就满足了这些条件。

本书从神经网络的基本概念出发，详细介绍了神经网络的学习过程、前向神经网络、后向反馈神经网络、循环神经网络、卷积神经网络、多层神经网络、感知神经网络等。在介绍神经网络基本知识的同时，应用当前最为广泛使用的数据分析软件 R，演示了相关的神经网络概念和模型的应用。同时，也介绍了当前最流行的神经网络模型的实现软件包，

例如 TensorFlow、Keras 等，并介绍如何把它们集成到 R 的应用中。

读者通过自己动手，可以更好地获得神经网络的实际运用体验，学习 R 的神经网络建模方法。本书从神经网络的基本概念入手，案例丰富，浅显易懂，无论是神经网络的初学者或者资深用户都能从本书受益。同时，读者也可以学习 R 语言的相关知识和操作。本书既可以用作神经网络或者人工智能等课程的教材，也可以用作实践工作的应用手册。

ABOUT THE AUTHOR

关于作者

Giuseppe Ciaburro 拥有 Università degli Studi di Napoli Federico II 的化学工程硕士学位，Seconda Università degli Studi di Napoli 的声学 and 噪声控制硕士学位。目前工作于 Università degli Studi della Campania “Luigi Vanvitelli” 的建筑环境控制实验室。

他有超过 15 年的 Python 和 R 语言编程经验，以及丰富的 MATLAB 使用经验，涉足氧化领域、声学 and 噪声控制领域。作为一名声学 and 噪声控制方面的专家，Giuseppe 有约 15 年的专业计算机课程授课经验，也制作了很多在线教学视频。同时，在专著、科学期刊论文和专题会议报告上都有贡献。他目前正在研究机器学习在声学 and 噪声控制中的应用。

Balaji Venkateswaran 是 AI 专家、数据科学家、机器学习实践者和数据库架构师。他在投资银行支付处理、电信计费和项目管理方面拥有 17 年以上的经验，也曾在 ADP、高盛、万事达和 Wipro 等大公司工作。Balaji 是数据科学、Hadoop 和 Tableau 方面的培训师。他拥有 Chennai 的 Great Lakes 管理学院商业分析专业硕士学位。

Balaji 拥有统计、分类、回归、模式识别、时间序列预测以及使用文本挖掘程序的非结构化数据相关的专业知识。他的主要兴趣是神经网络和深度学习。

Balaji 拥有 IBM SPSS、IBM Watson、IBM 大数据架构师、云架构师、CEH、Splunk、Salesforce、Agile CSM 和 AWS 等各种认证。

感谢我的父母和我生命中最重要的三个人——我的妻子 Aruna、儿子 Aadarsh 和 Abhitha。还要感谢 Packt 出版社的编辑，他们在本书的编写过程中给我提供了很多帮助。

ABOUT THE REVIEWER

关于审稿人

Juan Tomás Oliva Ramos 是墨西哥 Guanajuato 大学的环境工程师，拥有行政工程和质量专业硕士学位。他拥有 5 年的专利管理和开发、技术创新项目以及应用过程统计控制来开发技术解决方案的经验。自 2011 年以来，他一直扮演着统计学、创业和项目技术开发导师的角色。作为企业家导师，他在 Tecnológico Superior de Purísima del Rincón 开设了一个新的技术管理和创业部门。

Juan 是《Alfaomega》的评论员，曾撰写《*Wearable designs for Smart watches, Smart TVs and Android mobile devices*》。

他通过编程和自动化技术为改进某些操作开发了原型，并注册了专利。

非常感谢 Packt 出版社让我审阅这本优秀的书。感谢我的妻子 Brenda、女儿 Regina 和 Renata，以及即将诞生的 Angel Tadeo，感谢家人给予我的力量与快乐。

P R E F A C E

前言

神经网络是能有效解决复杂计算问题的最有吸引力的机器学习模型之一，用于解决人工智能（Artificial Intelligence, AI）和机器学习不同领域的各种问题。

本书解释了神经网络的优点，提供了高级主题相关的基础知识。本书从使用 `neuralnet` 包设计神经网络作为开始。然后介绍神经网络如何从数据中学习，以及背后的原理。本书涵盖了各种类型的神经网络，包括循环神经网络和卷积神经网络。通过本书，读者不仅可以学习如何训练神经网络，还可以探索这些神经网络的泛化。最后深入研究不同的神经网络模型，并与现实世界的用例相结合。

在本书的最后，读者将在实际案例的帮助下，学会在自己的应用程序中实现神经网络模型。

本书内容

第 1 章介绍人工神经网络（Artificial Neural Network, ANN）和人工智能的基本概念与理论，展示 ANN 和 AI 的简单应用程序与数学概念。还对 R 中的 ANN 函数进行介绍。

第 2 章介绍如何在图形模型中进行精确推断，并展示作为专家系统的应用程序。推断算法是学习和使用这类模型的基础。读者至少需要了解它们的用处及工作原理。

第 3 章阐述深度学习和深度学习中神经网络的使用。该章使用 R 的添加包介绍神经网络实现过程的细节，涵盖许多为深度学习设置的隐藏层，并使用实用的数据集来帮助读者

理解实现过程。

第 4 章介绍感知机以及使用它构建的应用程序，以及基于 R 的感知机实现。

第 5 章涵盖使用数据集训练神经网络的另一个案例，还通过使用函数 `plot()` 对输入层、隐藏层和输出层进行图形表示，帮助读者更好地理解神经网络。

第 6 章介绍循环神经网络和卷积神经网络及其在 R 中的实现。同时提出几个案例帮助读者了解基本概念。

第 7 章介绍不同领域的神经网络应用，以及神经网络如何在 AI 领域中使用。有助于读者了解神经网络算法的实际应用。读者可以采用不同的数据集、运行 R 代码来进一步增强自身的技能。

准备工作

本书着重于 R 环境中的神经网络，使用 R 3.4.1 和 RStudio 1.0.153 来建立各种应用程序、开源和企业级专业软件。本书专注于如何以最佳方式利用各种 R 添加包来构建现实世界的应用程序。本着这种精神，我们尽量保持所有代码的友好性和可读性。这将使读者能够轻松地读懂代码，并在不同的场景中轻松使用。

本书读者对象

本书适用于任何具有 R 和统计背景知识，同时希望使用神经网络从复杂数据中获得更好结果的人。如果你对人工智能和深度学习感兴趣，并希望提升自己，那么这本书就是你所需要的！

下载示例代码

读者可以从 <http://www.packtpub.com> 或者华章网站 <http://www.hzbook.com/> 下载本书的示例代码。

CONTENTS

目 录

译者序	
关于作者	
关于审稿人	
前言	
第 1 章 神经网络和人工智能概念1	
1.1 简介.....2	
1.2 神经网络的灵感.....3	
1.3 神经网络的工作原理.....4	
1.4 分层方法.....5	
1.5 权重和偏差.....6	
1.6 训练神经网络.....7	
1.6.1 有监督学习.....7	
1.6.2 无监督学习.....7	
1.7 epoch.....7	
1.8 激活函数.....8	
1.9 不同的激活函数.....8	
1.9.1 线性函数.....8	
1.9.2 单位阶跃激活函数.....9	
1.9.3 sigmoid 函数.....10	
1.9.4 双曲正切函数.....11	
1.9.5 线性修正单元函数.....11	
1.10 使用哪些激活函数.....12	
1.11 感知机和多层架构.....13	
1.12 前向和反向传播.....13	
1.13 逐步说明神经网络和激活函数.....14	
1.14 前馈和反馈网络.....16	
1.15 梯度下降.....17	
1.16 神经网络分类法.....17	
1.17 使用 R 语言神经网络添加包 neuralnet() 的简单示例.....19	
1.18 使用添加包 nnet() 进行实现.....24	
1.19 深度学习.....29	
1.20 神经网络的优缺点.....29	
1.21 神经网络实现的最佳实践.....30	
1.22 有关 GPU 处理的简要说明.....30	
1.23 小结.....31	
第 2 章 神经网络中的学习过程32	
2.1 机器学习.....33	
2.1.1 有监督学习.....34	

2.1.2	无监督学习	35
2.1.3	强化学习	36
2.2	训练和测试模型	37
2.3	数据循环	38
2.4	评估指标	39
2.5	学习神经网络	42
2.6	反向传播	43
2.7	神经网络学习算法的优化	45
2.8	神经网络中的有监督学习	46
2.8.1	波士顿数据集	46
2.8.2	对波士顿数据集进行神经网络回归	48
2.9	神经网络中的无监督学习	54
2.9.1	竞争学习	55
2.9.2	Kohonen SOM	57
2.10	小结	63

第 3 章 使用多层神经网络进行

	深度学习	64
3.1	DNN 简介	65
3.2	用于 DNN 的 R 语言	67
3.3	通过 neuralnet 建立多层神经网络	69
3.4	使用 H2O 对 DNN 进行训练和建模	78
3.5	使用 H2O 建立深度自动编码器	90
3.6	小结	91

第 4 章 感知神经网络建模——

基本模型

4.1	感知机及其应用	93
4.2	简单感知机——一个线性可分离分类器	96
4.3	线性分离	98
4.4	R 中的感知机函数	101
4.5	多层感知机	107
4.6	使用 RSNNs 在 R 中实现 MLP	108
4.7	小结	117

第 5 章 在 R 中训练和可视化

神经网络

5.1	使用神经网络进行数据拟合	120
5.1.1	探索性分析	124
5.1.2	神经网络模型	126
5.2	使用神经网络对乳腺癌进行分类	135
5.2.1	探索性分析	138
5.2.2	神经网络模型	143
5.2.3	网络训练阶段	146
5.2.4	测试神经网络	149
5.3	神经网络训练中的早期停止	152
5.4	避免模型中的过拟合	152
5.5	神经网络的泛化	154
5.6	神经网络模型中数据的缩放	154

5.7 集成神经网络来预测	155
5.8 小结	156

第 6 章 循环和卷积神经网络

6.1 循环神经网络	158
6.2 R 中的添加包 rnn	162
6.3 LSTM 模型	170
6.4 卷积神经网络	172
6.5 常见的 CNN 架构——LeNet	175
6.6 使用 RNN 进行湿度预测	176
6.7 小结	183

第 7 章 神经网络案例——

高级主题

7.1 TensorFlow 与 R 的集成	186
7.2 Keras 与 R 的集成	189
7.3 在 R 中使用 MNIST HWR	190
7.4 使用数据集 iris 建立 LSTM	199
7.5 使用自动编码器	203
7.6 使用 H2O 进行主成分分析	204
7.7 使用 H2O 建立自动编码器	207
7.8 使用添加包 darch 检测乳腺癌 ...	211
7.9 小结	217

神经网络和人工智能概念

在几个世纪以来进行的科学和哲学研究中，已经确定了作为人类智力基础的特殊机制。从这些机制的运作中获得灵感，我们可以创建能够模仿这些机制中的一部分的机器。问题是，这些机制还没有得到成功的模拟和整合，所以我们拥有的人工智能（Artificial Intelligence, AI）系统在很大程度上是不完整的。

改进这种机器的一个决定性的步骤是使用所谓的人工神经网络（Artificial Neural Network, ANN），从调节自然神经网络的机制开始，计划模拟人类思维。现在，软件可以模拟赢得国际象棋比赛所需的机制，或者根据语法规则将文本翻译成不同语言。

本章介绍 ANN 和 AI 的基本理论概念。读者需要对以下内容有基本的了解：

- 高中基础数学，各种微积分和函数（如 sigmoid 函数）知识。
- R 编程和 R 库的使用。

本章将介绍神经网络的基础知识，并尝试使用 R 建立一个模型。本章是神经网络和所有后续章节的基础。

本章涵盖：

- ANN 概念
- 神经元、感知机和多层神经网络

- 偏差、权重、激活函数和隐含层
- 前向和反向传播方法
- 图形处理单元的简介

在本章结束时，你将能够识别不同的神经网络算法和 R 提供的用来处理它们的工具。

1.1 简介

大脑是人体最重要的器官。它是我们执行所有功能的中央处理单元。大脑只有 1.5kg，拥有约 860 亿个神经元。神经元被定义为传递神经刺激或电化学信号的细胞。大脑是一个复杂的神经元网络，它通过几个相互关联的神经元系统来处理信息。了解大脑的功能一直是一项艰巨的任务。然而，由于计算技术的进步，现在可以人为编程神经网络。

ANN 的工作原理源于模拟同样正在试图解决问题的人脑的功能这一想法。相关常规方法的缺点以及 ANN 连续应用的问题已经在明确的技术环境中得到克服。

AI 或机器智能旨在将认知能力赋予计算机，通过编程来学习和解决问题。其目的是模拟具有人类智能的计算机。AI 不能完全模拟人的智慧；计算机只能通过程序来完成人脑某些方面的工作。

机器学习是 AI 的一个分支，它帮助计算机根据输入数据对自己进行编程。机器学习使 AI 能够基于数据解决问题。ANN 是机器学习算法的一个例子。

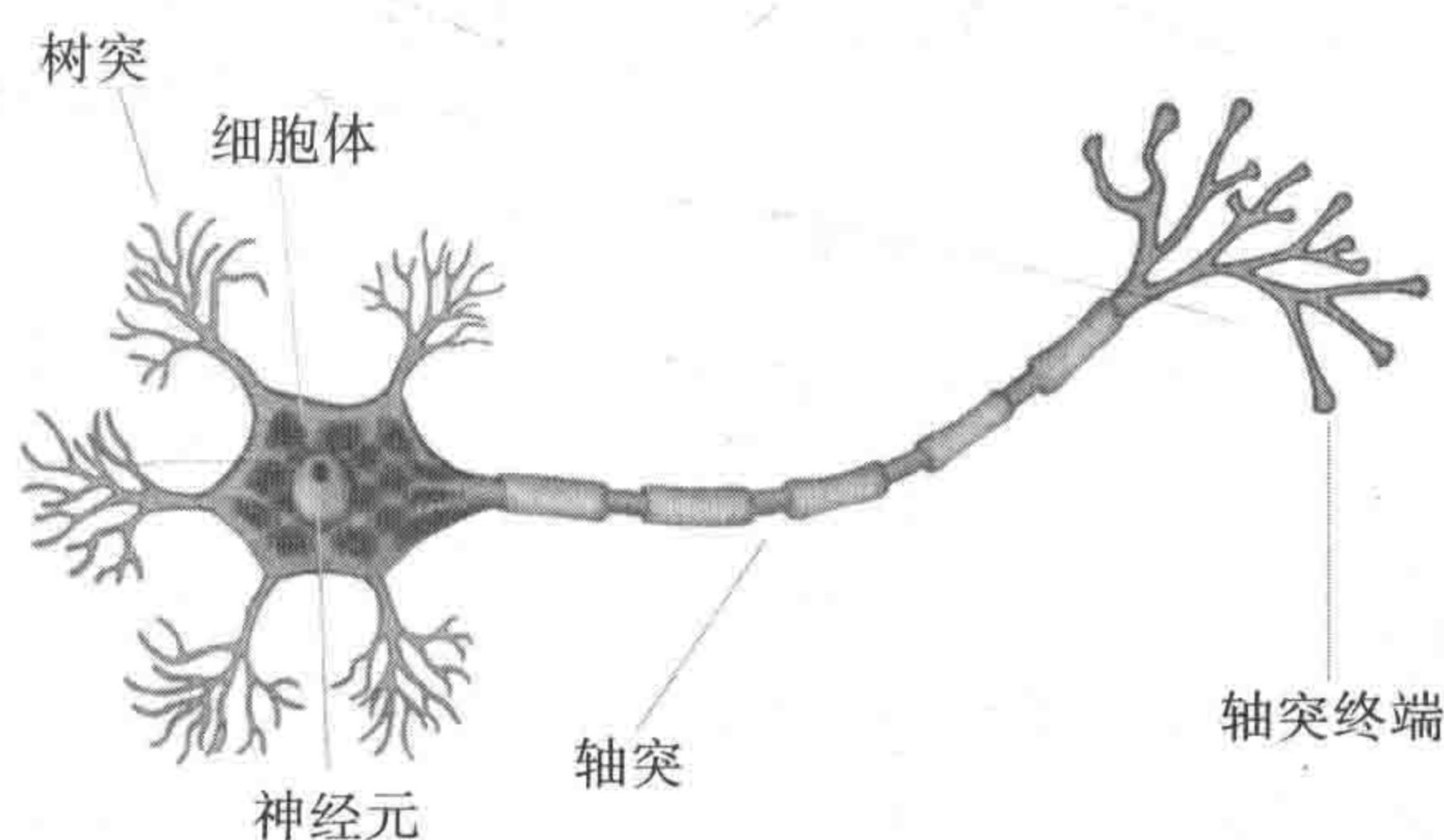
深度学习（Deep learning, DL）是具有更多处理层的复杂神经网络集，开发过程高度抽象。DL 通常用于复杂的任务，如图像识别、图像分类和手写识别。

大多数读者认为神经网络难以学习并将其当作黑匣子。本书打算打开这个黑匣子，帮助读者学习神经网络在 R 中的内部实现。通过知识管理，读者可以看到许多，神经网络发挥重要作用的场景，见下图。



1.2 神经网络的灵感

神经网络的灵感来源于人脑的工作方式。人类的大脑可以使用人类感知（特别是视觉）发送的数据来处理大量的信息。该处理由神经元完成，神经元对通过它们的电信号起作用，并施加触发器逻辑，例如打开和关闭门以使信号通过。下图显示了神经元的结构。



每个神经元的主要成分是：

- 树突 (dendrite)：每个神经元的入口点，信号以电脉冲的形式从神经网络中的其他神经元接受输入。
- 细胞体 (cell body)：它从树突输入端产生推断并决定采取什么行动。
- 轴突终端 (axon terminal)：以电脉冲的形式将输出传输到下一个神经元。

每个神经元只有在信号超过某个阈值时才处理它。神经元要么放电要么不放电；要么为 0 要么 1。

AI 一直出现科幻电影和小说里。AI 中的 ANN 自 20 世纪 50 年代就已经存在，但由于计算架构和性能的提高，它在过去 10 年中变得更加重要。计算处理方面的重大进展推动了以下技术的发展：

- 大规模并行计算。
- 分布式表示和计算。
- 学习和泛化能力。
- 容错。
- 能耗低。

在数值计算和符号操作的领域，对于解决集中式架构的问题，现代计算机在更大程度上超越了人类。然而，在模式识别、降噪和优化领域却落后于人类。一个小孩可以在巨大的人群中找出他 / 她的妈妈，但是具有集中式架构的计算机无法做到这一点。

这也是大脑的生物神经网络超越机器的地方，因此人们想开发一种模仿人类大脑的松散分布式替代架构。

ANN 是由大量具有许多互连的简单处理器组成的大规模并行计算系统。

全球领先的通讯社之一 Guardian，通过上传所有档案的快照，利用大数据来存档。然而，用户复制内容并在其他地方使用是受限制的。为了克服这一缺点，人们可以使用 ANN 进行文本模式识别，将图像转换为文本文件，然后根据最终用户的需要再将其转换为任何格式。

1.3 神经网络的工作原理

类似于生物神经元结构，ANN 将神经元定义为中央处理单元，其执行数学运算以从一组输入生成一个输出。神经元的输出是输入的加权和加上偏差的函数。如果接收到的信号总量超过激活阈值，则每个神经元都执行非常简单的操作，如下图所示。