HZ BOOKS 华章IT

[PACKT] PUBLISHING

智能系统与技术丛书

Neural Network Programming with Java

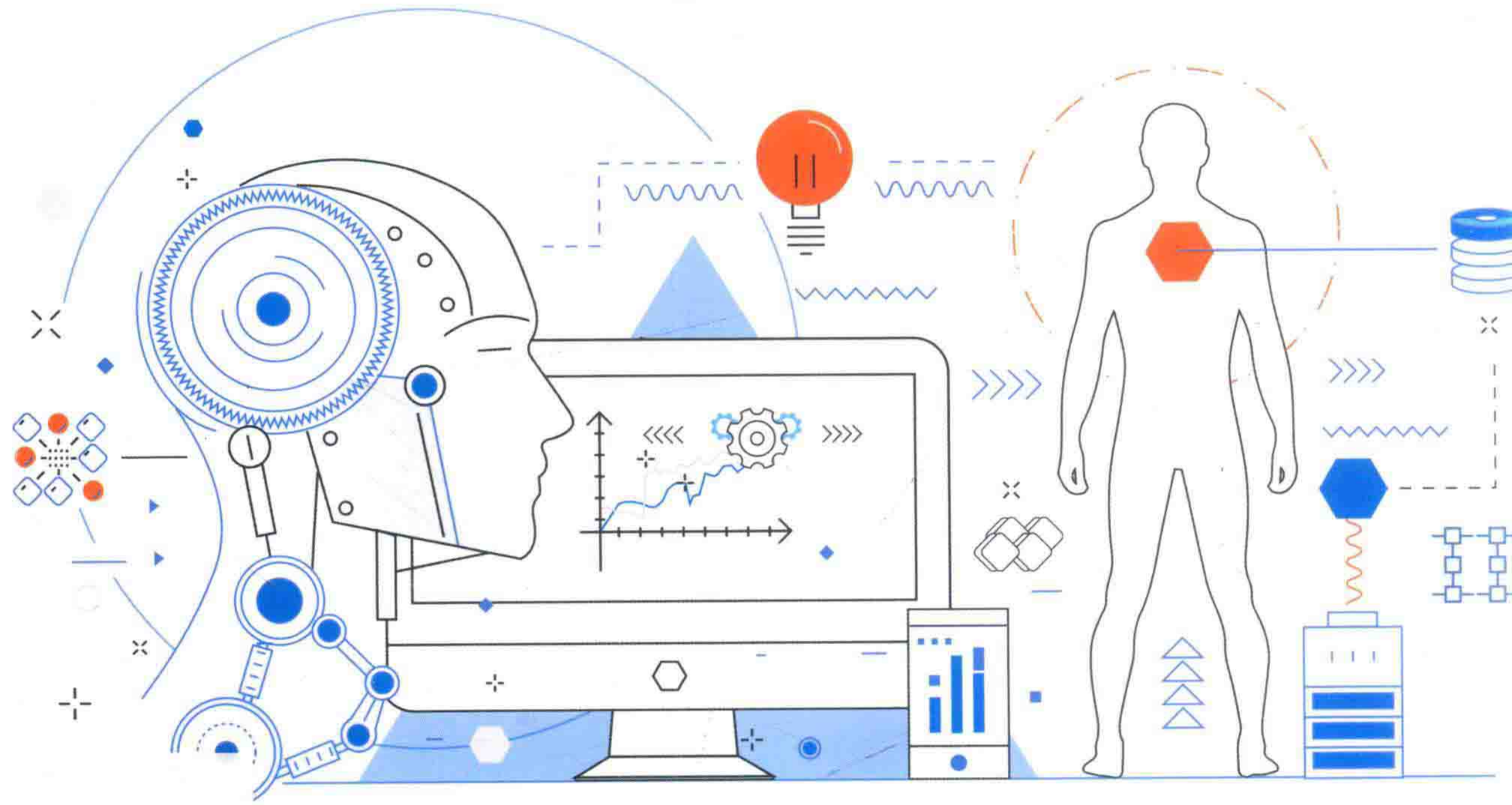
Second Edition

神经网络编程实战

Java语言实现

（原书第2版）

[巴西] 法比奥 · M. 索尔斯（Fabio M. Soares）
艾伦 · M. F. 索萨（Alan M. F. Souza） 著
王彩霞 夏妍 译



机械工业出版社
China Machine Press

Neural Network Programming with Java
Second Edition

神经网络编程实战
Java语言实现
（原书第2版）

[巴西] 法比奥·M. 索尔斯（Fabio M. Soares）
艾伦· M. F. 索萨（Alan M. F. Souza） 著
王彩霞 夏妍 译

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

神经网络编程实战：Java 语言实现（原书第 2 版）/（巴西）法比奥·M. 索尔斯（Fabio M. Soares），（巴西）艾伦·M. F. 索萨（Alan M. F. Souza）著；王彩霞，夏妍译．—北京：机械工业出版社，2018.5
（智能系统与技术丛书）
书名原文：Neural Network Programming with Java, Second Edition

ISBN 978-7-111-60012-1

I. 神… II. ①法… ②艾… ③王… ④夏… III. ①人工神经网络 -JAVA 语言 程序设计 IV. ①TP183 ②TP312.8

中国版本图书馆 CIP 数据核字（2018）第 109558 号

本书版权登记号：图字 01-2017-7506

Fabio M. Soares, Alan M. F. Souza : *Neural Network Programming with Java*, Second Edition (ISBN: 978-1-78712-605-3).

Copyright © 2017 Packt Publishing. First published in the English language under the title "Neural Network Programming with Java, Second Edition".

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。

神经网络编程实战：Java 语言实现（原书第 2 版）

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：谢晓芳　　责任校对：殷 虹
印　　刷：三河市宏图印务有限公司　　版　　次：2018 年 6 月第 1 版第 1 次印刷
开　　本：186mm×240mm 1/16　　印　　张：14
书　　号：ISBN 978-7-111-60012-1　　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

众所周知，人工智能是当下非常热门的技术。其概念在 20 世纪 80 年代就已经炒得火热，但是由于软硬件两方面的技术局限使其两度陷入低谷。而如今，各大 IT 巨头（诸如谷歌、Facebook、Apple、百度、腾讯、阿里等）纷纷在人工智能领域密集布局，通过巨额研发投入、人才储备、投资并购及开源合作等方式极力打造各自的人工智能生态圈。人工智能已经成为这个时代激动人心、值得期待的技术，或将成为未来 10 年乃至更长时间内 IT 产业发展的焦点。究其原因，除了互联网大数据积累、计算能力提升及计算成本降低等推动因素外，其本质原因是孜孜不倦积累 30 多年的神经网络技术的集中爆发。

神经网络（Neural Network）是人工智能领域最重要的基础模型之一，也是目前非常热门的研究方向——深度学习的核心。作为一门重要的机器学习技术，其已经广泛应用于自动驾驶、机器翻译、语音识别、图像识别等许多前沿领域。而当前人工智能领域的诸多应用平台（如智能客服等）都基于 Java 服务框架开发，因此，通过 Java 实现机器学习算法对于业务平台实现真正的智能化非常重要。

遗憾的是，现有的书籍要么只阐述机器学习方面的算法理论，要么只阐述 Java 理论，缺少基于工程性语言的机器学习算法实现，由此带来的问题是智能化应用产品的落地和实现过程的复杂化。

作为一本集神经网络理论、神经网络 Java 实现及优秀案例于一体的书，本书以一种简单易懂、循序渐进的方式介绍神经网络。本书由 10 章构成。首先系统、全面

地阐述了神经网络的相关概念、知识点及特征。然后重点介绍了神经网络学习过程的细节，如何用 Java 实现神经网络特性及设计神经网络架构，如何优化调整神经网络参数等。最后介绍了一些经典案例。结合 Java 实现神经网络架构是本书的一个主要特色。这是一本基于 Java 语言阐述神经网络架构特性、设计实现过程又包含经典实践案例的优秀书籍。

本书的读者对象如下：1）神经网络初学者，本书全面阐述了神经网络的相关概念、知识点以及神经网络架构的创建过程；2）有 Java 背景且对智能化方向感兴趣的软件开发者，本书介绍了如何基于 Java 实现神经网络核心算法；3）智能化软件开发人员，本书提供了基于神经网络 Java 实现的诸多案例，为解决实践中的问题提供了很好的借鉴。本书不需要读者有深厚的数学基础，但是懂一些机器学习基础知识更有助于理解本书。

参与本书翻译工作的人员主要有：百度研发工程师王彩霞和蔚来汽车研发工程师夏妍。非常感谢百度地图刘懿的牵线引荐，让我能有机会和机械工业出版社合作。衷心感谢刘懿、百度美国大数据实验室的杨涛、百度推荐技术平台部的尚斌及网易产品经理马庆提出的修改建议。也非常感谢清华大学公派加州大学伯克利分校博士顾维玺、机械工业出版社张梦玲编辑的无私帮助。

限于水平，对中文语言的表达及内容的理解难免存在不当之处，在此敬请读者批评指正。本人十分荣幸能有机会成为本书的译者，认真阅读本书也一定会使你受益匪浅。

王彩霞

2018 年 5 月于厦门

此为试读，需要完整PDF请访问：www.ertongbook.com

ABOUT THE AUTHOR

作者和审校者简介

作者简介

Fabio M. Soares 目前是位于巴西北部的帕拉联邦大学（UFPA）的博士研究生。他几乎对所有领域的技术都充满热情，2004 年起开始设计神经网络解决方案，此后将该技术应用于电信、工业过程控制和建模、水力发电、财务应用、零售客户分析等多个领域。他的研究课题涉及针对数据驱动模型的监督学习。2017 年以来，他主要在铝冶炼和铁镍合金方面进行化学过程建模及控制的研究，同时也作为计算机编程和人工智能课程的讲师，从事教学工作。作为一名活跃的研究者，他已经在诸多会议和期刊上发表了数篇论文，参与 4 本书的编写。

Alan M. F. Souza 是 IESAM（Instituto de Estudos Superiores da Amazonia）的一名 IT 工程师，他拥有帕拉联邦大学软件项目管理专业的研究生学历和工业过程专业（应用计算）的硕士学位。2009 年以来，他一直从事神经网络相关工作，并且自 2006 年起，他和巴西 IT 公司共同致力于 Java、PHP、SQL 及其他编程语言的相关研发工作。他非常热爱编程和计算智能。目前他是 UNAMA 的教授，同时也是 UFPA 的博士研究生。

审校者简介

Charles Griffiths 是一名对技术、医学、经济学和营养学感兴趣的软件工程师和大学教授。他在 GitHub 和 Unity Asset Store 上都发布过源码。

无论走到哪里，都很感谢那些指引我前进的朋友，感谢他们一直以来对我的支持和信任。

PREFACE

前　言

程序员需要持续不断地学习，而且经常会面对新技术和新方法的挑战。生活中人们虽然习惯了重复的事情，但也会经历新的事情。学习过程是科学界最有趣的话题之一，很多尝试都试图描述或者再现人类的学习过程。

本书的主要挑战是学习并掌握业界最新的内容。虽然神经网络这个名字可能看起来很奇怪，甚至可能误认为它是关于神经学的，但是我们通过把重点放在你决定购买这本书的原因上来简化这些细微差别。我们打算建立一个框架，告诉你神经网络其实很简单，很容易理解，你不需要有足够的先验知识，就完全可以理解本书提到的概念。

因此，我们希望你充分掌握本书的内容，在面对棘手问题时，能始终以初学者的态度运用神经网络的功能来解决。本书对提到的每个概念都用简单的语言进行解释，但理解它也需要一定的技术背景。本书的目的是让你了解智能应用可以通过简单语言编写。

各章概览

第 1 章主要介绍神经网络的概念，解释基本神经元结构（单层感知机、学习机），以及激活函数、权重和学习算法。此外，该章还演示了用 Java 创建基本神经网络的整个过程。

第 2 章主要介绍神经网络学习过程的细节，解释几个有用的概念，如训练、测试和验证，演示如何实现训练和验证算法、如何进行误差评估。

第 3 章主要讨论感知机和监督学习的特性，展示这类神经网络的训练算法，以及如何用 Java 实现这些特性。

第 4 章主要介绍无监督学习和自组织映射，即 Kohonen 神经网络在分类和聚类问题中的应用。

第 5 章主要阐述如何用神经网络解决天气预报的问题，你会看到来自不同地区、不同时间的历史天气数据记录，并学习如何在神经网络训练之前对数据进行预处理。

第 6 章主要介绍分类问题，这属于监督学习的范畴。运用患者的数据构建基于神经网络的专家系统，专家系统能够根据患者的症状给出诊断结果。

第 7 章讨论如何应用无监督学习算法和神经网络实现聚类，进而实现客户画像聚类。

第 8 章主要介绍另一种涉及神经网络的常见任务：光学字符识别（OCR）。OCR 非常有用，它显示了神经网络强大的学习能力。

第 9 章主要介绍神经网络优化的相关技术，如输入选择，切分训练数据集、验证数据集和测试数据集的较优方法，以及数据过滤和隐含神经元个数的选择。

第 10 章主要介绍神经网络领域的新技术动态，启发你理解并设计出适用于更复杂问题的新策略。

附录内容为在线内容，可以通过以下链接下载：https://www.packtpub.com/sites/default/files/downloads/Neural_Network_Programming_with_Java_SecondEdition_Appendices.pdf。附录内容主要涉及搭建 Netbeans 开发环境的详细步骤，搭建 Eclipse 开发环境的详细步骤。

阅读准备

需要 Netbeans（www.netbeans.org）或者 Eclipse（www.eclipse.org）软件，两者都是免费的，可以从官方网站下载。

本书读者对象

本书适合以下 Java 开发者：想知道如何运用神经网络的功能开发更智能的应用。同时本书适用于那些处理大量复杂数据并希望在日常应用中有更高效率的人士。本书的读者最好具有一些统计计算的基础知识。

下载示例代码

可以从网站 http://www.packtpub.com 或华章网站 www.hzbook.com 下载本书中的示例代码。

CONTENTS

目　录

译者序

作者和审校者简介

前言

第 1 章 神经网络入门 …… 1

1.1 探索神经网络 …… 1

1.2 人工神经网络 …… 2

1.2.1 神经网络是如何组织的 …… 3

1.2.2 基本元素——人工神经元 …… 3

1.2.3 赋予神经元生命——激活函数 …… 4

1.2.4 可变参数——权重 …… 5

1.2.5 额外参数——偏置 …… 6

1.2.6 由部分到整体——层 …… 6

1.2.7 神经网络体系结构 …… 7

1.2.8 单层网络 …… 7

1.2.9 多层网络 …… 8

1.2.10 前馈网络 …… 8

1.2.11 反馈网络 …… 8

1.3 从无知到认知——学习过程 …… 9

1.4 开始编程——神经网络实践 …… 10

1.5 神经元类 …… 12

1.6 NeuralLayer 类 …… 14

1.7 ActivationFunction 接口 …… 15

1.8 神经网络类 …… 15

1.9 运行程序 …… 17

1.10 本章小结 …… 19

第 2 章

神经网络学习 …… 20

2.1 神经网络的学习能力 …… 21

2.2 学习模式 …… 22

2.2.1 监督学习 …… 22

2.2.2 无监督学习 …… 22

2.3 学习过程 …… 23

2.3.1 寻找损失函数最优下降方向 …… 24

2.3.2 在学习过程中更新权重 …… 25

2.3.3 计算损失函数 …… 25
2.3.4 一般误差和总体误差 …… 27
2.3.5 神经网络的迭代学习什么时候停止比较好 …… 27
2.4 学习算法示例 …… 28
2.4.1 δ规则 …… 29
2.4.2 学习率 …… 30
2.4.3 实现δ规则 …… 30
2.4.4 δ规则学习的核心——train 和 calcNewWeight 方法 …… 31
2.4.5 另一种学习算法——Hebbian 学习 …… 34
2.4.6 学习机 …… 35
2.5 在实践中理解学习过程 …… 37
2.6 测试 …… 41
2.7 本章小结 …… 43
第 3 章 感知机和监督学习 …… 44
3.1 监督学习——训练神经网络 …… 45
3.1.1 分类——寻找合适的类别 …… 45
3.1.2 回归——将实际输入映射到输出 …… 46
3.2 一个基本的神经结构——感知机 …… 48
3.2.1 应用和限制 …… 49
3.2.2 线性可分 …… 49
3.2.3 “异或”问题 …… 50
3.3 多层感知机 …… 52
3.3.1 MLP 属性 …… 52
3.3.2 MLP 权重 …… 53
3.3.3 递归 MLP …… 54
3.3.4 编码实现 MLP …… 54
3.4 MLP 学习 …… 55
3.4.1 反向传播算法 …… 56
3.4.2 动量项 …… 58
3.4.3 编码实现反向传播 …… 58
3.4.4 Levenberg-Marquardt 算法 …… 62
3.4.5 编码实现基于矩阵代数的 Levenberg-Marquardt 算法 …… 64
3.4.6 极限学习机 …… 66
3.5 实例 1——基于δ规则和反向传播的“异或”问题 …… 69
3.6 实例 2——预测入学状态 …… 72
3.7 本章小结 …… 75
第 4 章 自组织映射 …… 76
4.1 无监督神经网络 …… 76
4.2 无监督学习算法 …… 77
4.2.1 竞争学习 …… 78
4.2.2 竞争层 …… 80
4.3 Kohonen 自组织映射 …… 82
4.3.1 将神经网络代码扩展至 Kohonen …… 83
4.3.2 零维 SOM …… 84
4.3.3 一维 SOM …… 84
4.3.4 二维 SOM …… 85

此为试读，需要完整PDF请访问：www.ertongbook.com

4.3.5 2D 竞争层 …… 87
4.3.6 SOM 学习算法 …… 89
4.3.7 邻近神经元的影响——邻域函数 …… 90
4.3.8 学习率 …… 91
4.3.9 竞争学习的一个新类 …… 92
4.3.10 SOM 可视化 …… 95
4.3.11 绘制训练数据集和神经元权重的 2D 图 …… 97
4.3.12 测试 Kohonen 学习 …… 99
4.4 本章小结 …… 105
第 5 章 预报天气 …… 106
5.1 神经网络用于回归问题 …… 106
5.2 加载 / 选择数据 …… 108
5.2.1 创建辅助类 …… 108
5.2.2 从 CSV 文件加载数据集 …… 111
5.2.3 创建时序结构 …… 112
5.2.4 丢弃 NaN …… 113
5.2.5 获取天气数据 …… 114
5.2.6 天气变量 …… 115
5.3 选择输入和输出变量 …… 115
5.4 预处理 …… 117
5.4.1 归一化 …… 117
5.4.2 应用 NeuralDataSet 处理归一化 …… 121
5.4.3 应用学习算法进行归一化 …… 123
5.4.4 天气预报的 Java 实现 …… 123
5.4.5 收集天气数据 …… 123
5.4.6 延迟变量 …… 126
5.4.7 加载数据并开始运行 …… 126
5.4.8 相关性分析 …… 128
5.4.9 创建神经网络 …… 131
5.4.10 训练和测试 …… 131
5.4.11 可视化神经网络的输出 …… 133
5.5 神经网络实验设计 …… 134
5.5.1 设计实验 …… 134
5.5.2 结果和模拟 …… 135
5.6 本章小结 …… 138
第 6 章 疾病分类识别 …… 139
6.1 分类问题的基础 …… 139
6.1.1 分类数据 …… 140
6.1.2 处理分类数据 …… 141
6.2 逻辑回归 …… 142
6.2.1 多分类与二分类 …… 143
6.2.2 混淆矩阵 …… 144
6.2.3 敏感性与特异性 …… 144
6.2.4 实现混淆矩阵 …… 145
6.3 分类神经网络 …… 147
6.4 用神经网络进行疾病识别 …… 147
6.4.1 乳腺癌识别 …… 148
6.4.2 糖尿病识别 …… 151
6.5 本章小结 …… 154
第 7 章 客户画像聚类 …… 155
7.1 聚类任务 …… 156

7.1.1 聚类分析 …… 156
7.1.2 聚类评估和验证 …… 157
7.1.3 实现 …… 158
7.1.4 外部验证 …… 159
7.2 应用无监督学习 …… 159
7.3 画像过程 …… 160
7.3.1 预处理 …… 160
7.3.2 Java 实现 …… 161
7.3.3 信用卡——客户画像信用分析 …… 161
7.3.4 产品画像 …… 165
7.3.5 多少个簇合适 …… 166
7.4 本章小结 …… 167

第 8 章 文本识别 …… 168
8.1 模式识别 …… 168
8.1.1 类已知 …… 169
8.1.2 类未知 …… 170
8.2 神经网络用于模式识别 …… 171
8.2.1 数据预处理 …… 171
8.2.2 文本识别（光学字符识别）…… 172
8.2.3 数字识别 …… 172
8.2.4 数字表示 …… 172
8.2.5 Java 实现 …… 173
8.2.6 数据生成 …… 173
8.2.7 神经结构 …… 174
8.2.8 实验 …… 174
8.2.9 结果 …… 176
8.3 本章小结 …… 179

第 9 章 神经网络优化与调整 … 180
9.1 神经网络实现的常见问题 …… 181
9.2 输入数据选择 …… 181
9.2.1 数据相关性 …… 182
9.2.2 数据转换 …… 183
9.2.3 降维 …… 183
9.2.4 数据过滤 …… 184
9.2.5 交叉验证 …… 186
9.2.6 神经网络结构选择 …… 187
9.3 在线重训练 …… 189
9.3.1 随机在线学习 …… 190
9.3.2 实现 …… 190
9.3.3 应用 …… 191
9.4 自适应神经网络 …… 193
9.4.1 自适应谐振理论 …… 193
9.4.2 实现 …… 194
9.5 本章小结 …… 195

第 10 章 神经网络当前趋势 … 196
10.1 深度学习 …… 196
10.2 深度架构 …… 198
10.2.1 如何用 Java 实现深度学习 …… 199
10.2.2 神经模糊 …… 201
10.2.3 神经遗传 …… 203
10.3 实现混合神经网络 …… 204
10.4 本章小结 …… 207

参考文献 …… 208

CHAPTER 1

第1章

神经网络入门

本章将介绍神经网络及其设计目的。作为后续章节的基础，本章主要介绍神经网络的基本概念。本章将讨论以下主题：

- 人工神经元
- 权重和偏置
- 激活函数
- 神经元层
- 使用 Java 实现神经网络

1.1 探索神经网络

听到神经网络这个词，从直觉上我们会想到大脑，的确，我们可以将大脑看成一个大型的天然神经网络。然而，人工神经网络又是什么呢？人工是一个与天然相对的词，我们首先想到的就是人工大脑或者机器人，这就是所谓的人工。在这种情况下，受人脑的启发，我们创建出一个和人脑相似的结构，称之为人工智能。

ANN 初学者可能认为本书是讲如何构建智能系统的，例如人工大脑，智能系统能用 Java 代码模拟人类思维，是这样吗？答案是肯定的，但是，我们不会像电影《黑客帝国》中那样讨论如何创造人工思考机器；而会介绍人工神经网络解决方案的

设计过程，它能够利用整个 Java 框架，从原始数据抽象知识。

1.2　人工神经网络

在不了解神经网络的起源和相关术语的情况下，无法讨论神经网络。本书中神经网络（NN）和人工神经网络（ANN）是同义词，尽管 NN 因涵盖自然神经网络而更加通用。那么，什么是 ANN 呢？下面探究这个词的历史。

20 世纪 40 年代，神经生理学家 Warren McCulloch 和数学家 Walter Pitts 将神经学基础和数学运算结合起来，设计了人工神经元的第一个数学实现。当时，人脑被大量研究以弄懂那些潜在及神秘的行为，不过主要在神经学领域。众所周知，生物神经元结构有一个细胞核和多个树突（接收来自其他神经元传入的信号），以及一个轴突（将信号传递给其他神经元），如图 1-1 所示。

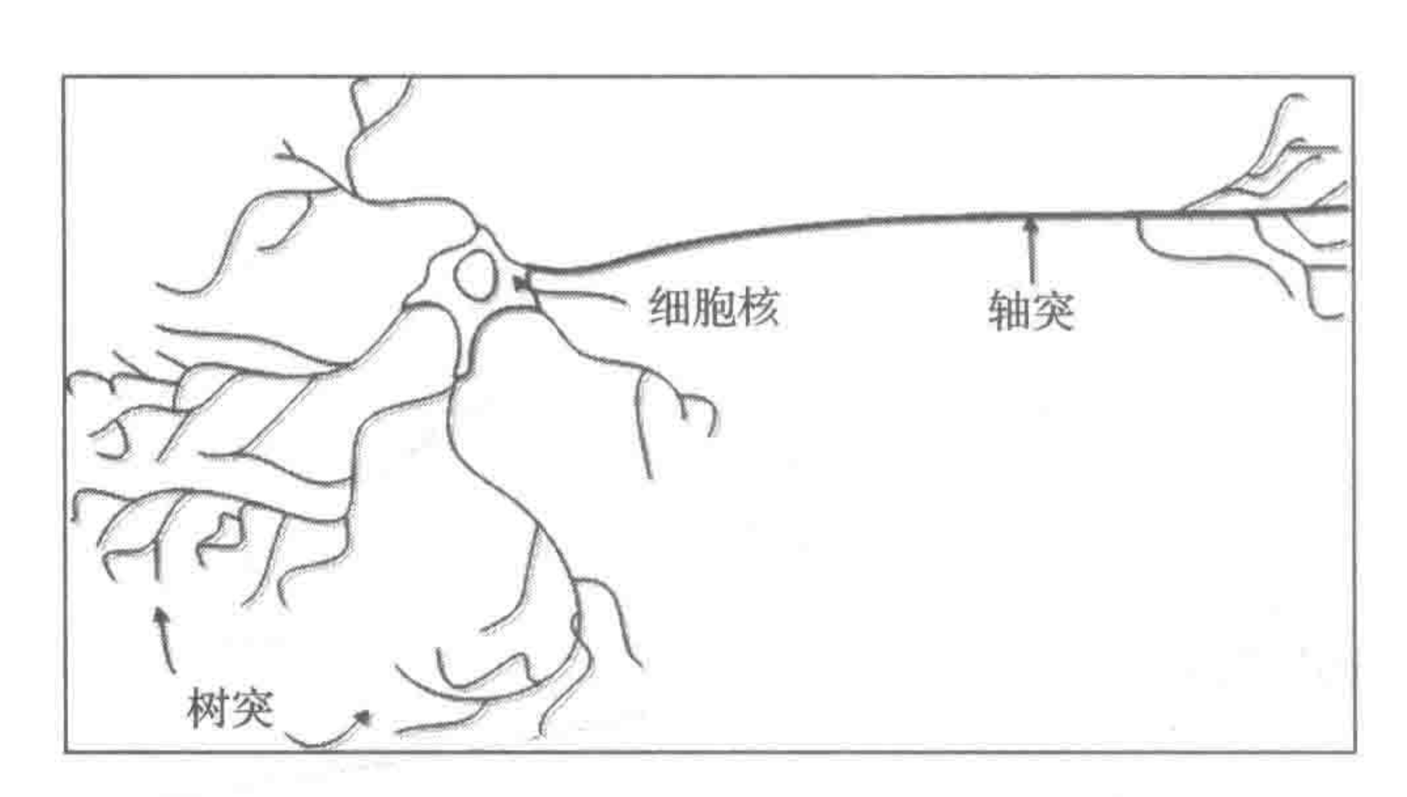


图 1-1　生物神经元结构

McCulloch 和 Pitts 的创新是在神经元模型中加入了数学成分，他们假设神经元是一个简单的处理器，用于合并所有输入信号并产生新的信号以激活其他神经元，如图 1-2 所示。

此外，考虑到大脑由数十亿个神经元组成，每个都与成千上万个其他神经元相联系，产生了数万亿的连接，因此我们讨论的是一个巨大的网络结构。基于这个事实，McCulloch 和 Pitts 为单个神经元设计一个简单的模型，最初用来模拟人类视觉。当

时可用的计算器或计算机资源虽然有限，但能很好地处理数学运算。即使今天，像视觉和声音识别这类任务，没有特殊的框架仍然很难编程实现。然而，相比于复杂的数学运算，人脑可以更高效地识别声音和图像，这激起了科学家和研究人员的兴趣。

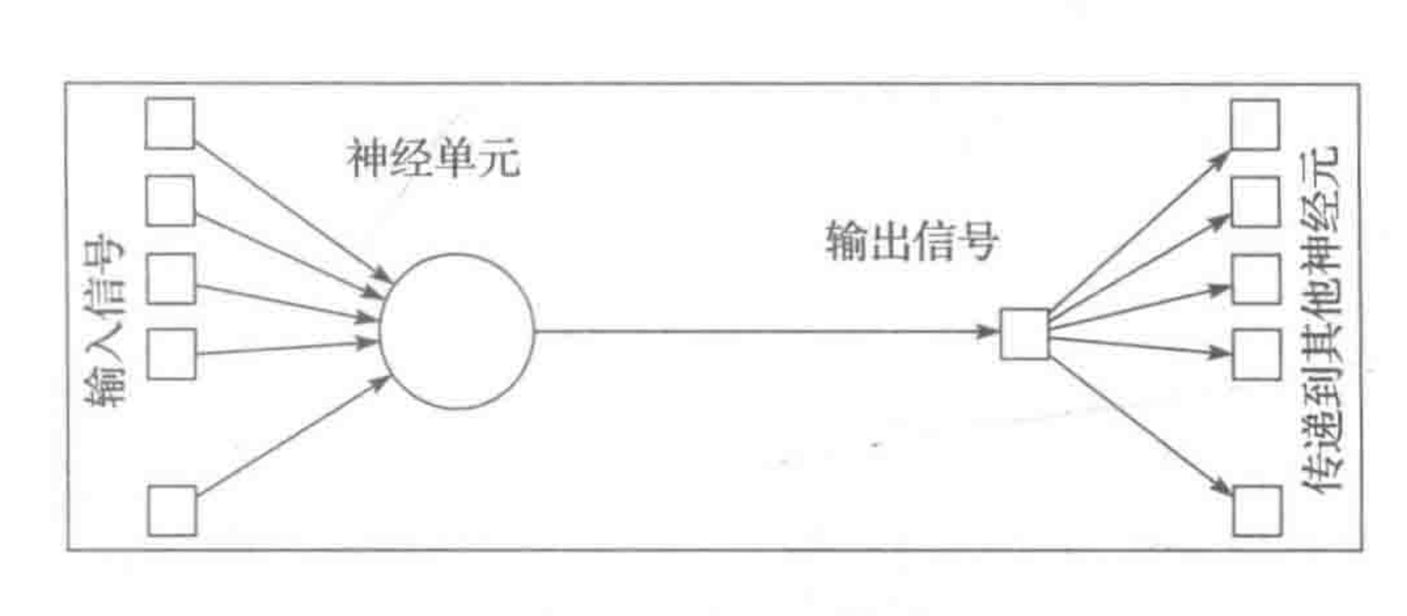


图 1-2　神经元模型

然而，众所周知，人脑执行的所有复杂活动都基于所学知识，为了克服传统算法在人类易于解决的问题上所面临的困难，我们设计了 ANN 解决方案，期望它具备基于外部刺激（数据），通过自学习来解决问题的能力（见表 1-1）。

表 1-1　人类和计算机可解决的任务

人类可解决的任务	计算机可解决的任务
图像分类	复杂计算
语音识别	语法纠错
人脸辨识	信号处理
基于经验预测事件	操作系统管理

1.2.1　神经网络是如何组织的

结合人脑的特点和结构，可以说 ANN 是一种自然启发的方法。每个神经元与许多其他神经元相接，这些神经元又会和其他大量神经元相连，形成一个高度互连的结构。本书后续章节将会介绍，神经元之间的连通性解释了学习能力，因为每个连接都可以根据刺激和期望目标进行配置。

1.2.2　基本元素——人工神经元

我们来看看最基本的人工神经元素——人工神经元。已证明生物神经元是信号处

理器，神经元中的树突会根据接收到信号的强度和振幅，发送信号到轴突。可以这样认为，神经元在输入上有一个信号收集器，在输出上有一个激活单元，它可以触发一个新的信号，然后传递给其他神经元，如图 1-3 所示。

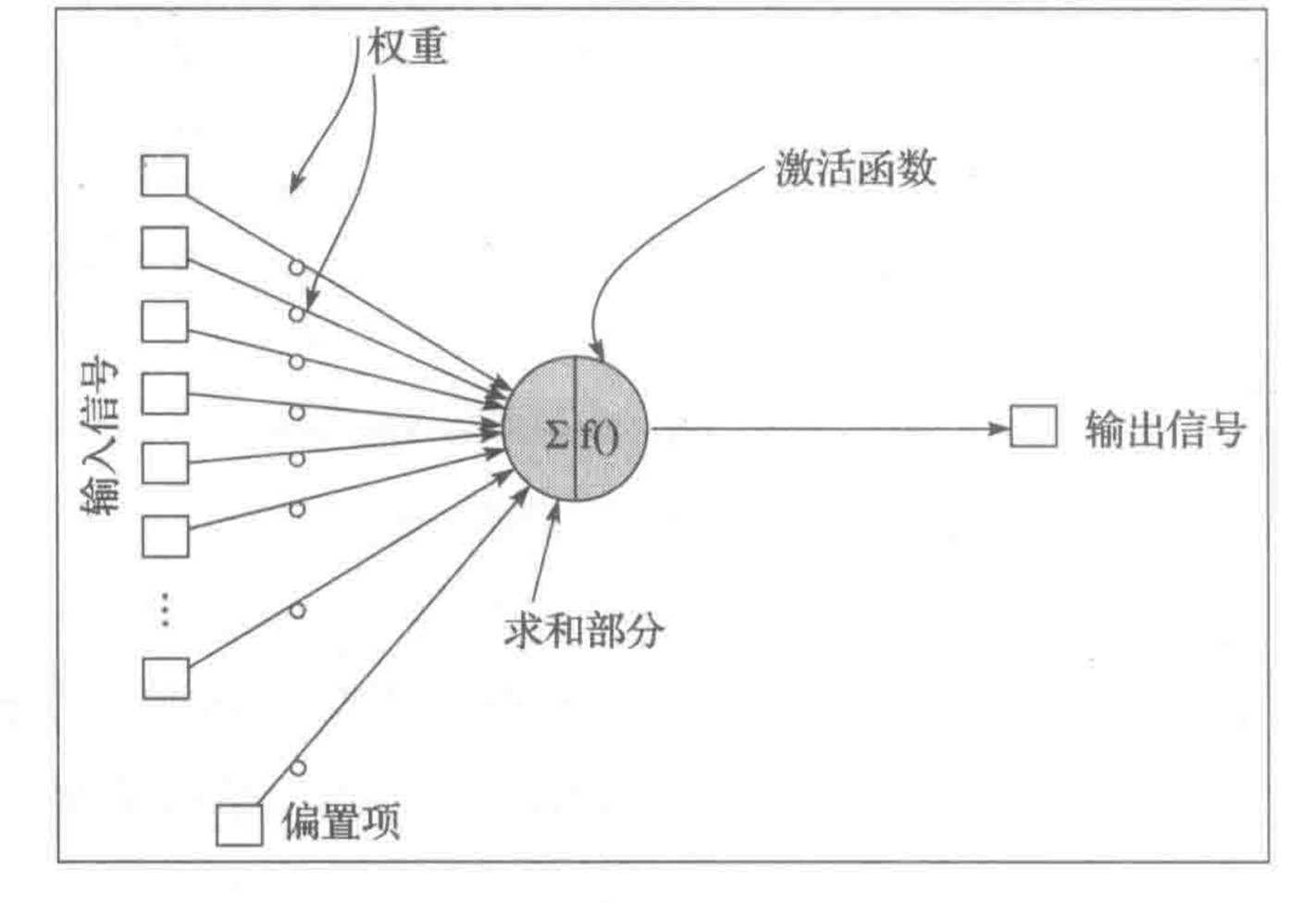


图 1-3　人工神经元

> **提示**　生物神经元中，存在一个潜在的阈值，一旦达到该值，会触发轴突并传递信号到其他神经元。这一行为可以通过激活函数来模拟，已经证明，激活函数在表示神经元的非线性行为方面非常有用。

1.2.3　赋予神经元生命——激活函数

激活函数是指一个神经元根据输入信号，执行计算并产生输出。从数学方面讲，激活函数用于为神经网络模型的处理加入非线性因素，从而提供人工神经网络的非线性行为，这对模拟生物神经元的非线性特征非常有用。激活函数通常是一个非线性函数，输出限制在某个区间范围内，但某些特定情况下，也可以是线性函数。

虽然任何函数都可以用作激活函数，但是本章主要介绍常用的几种，见表 1-2。

此为试读，需要完整PDF请访问：www.ertongbook.com