

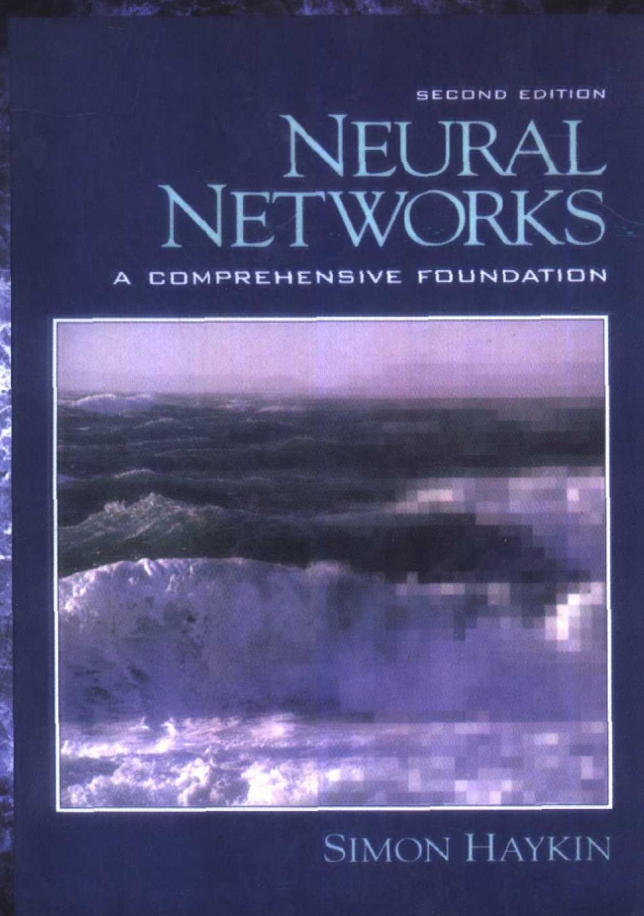
原书第2版

PEARSON
Prentice
Hall

计 算 机 科 学 丛 书

神经网络原理

Simon Haykin 著 叶世伟 史忠植 译



Neural Networks

A Comprehensive Foundation, Second Edition



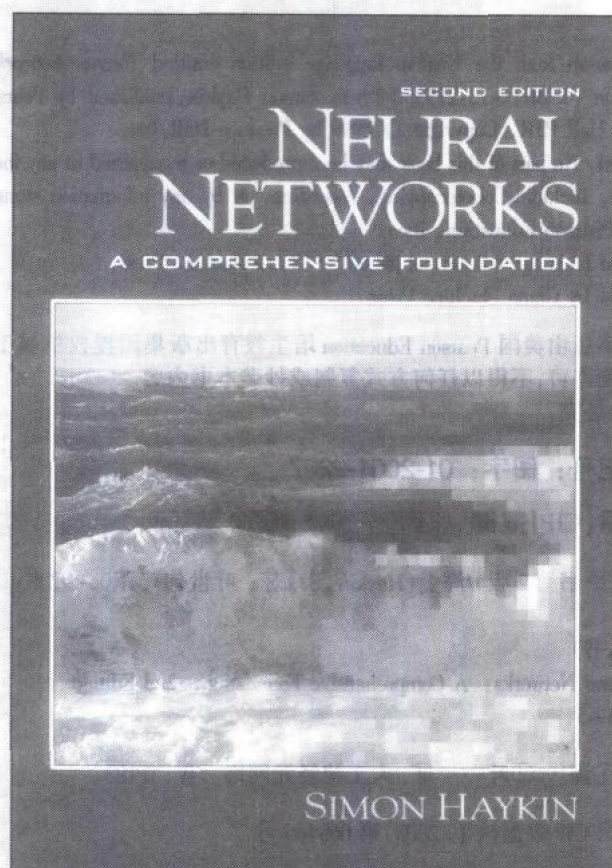
机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

原书第2版

神经网络原理

Simon Haykin 著 叶世伟 史忠植 译



Neural Networks
A Comprehensive Foundation, Second Edition



机械工业出版社
China Machine Press

神经网络是计算智能和机器学习研究的最活跃的分支之一。本书全面系统地介绍神经网络的基本概念、系统理论和实际应用。

本书包含四个组成部分：导论，监督学习，无监督学习，神经网络动力学模型。导论部分介绍神经元模型、神经网络结构和机器学习的基本概念和理论。监督学习讨论感知机学习规则，有监督的 Hebb 学习，Widrow-Hoff 学习算法，反向传播算法及其变形，RBF 网络，正则化网络，支持向量机以及委员会机器。无监督学习包括主分量分析，自组织特征映射模型的竞争学习形式，无监督学习的信息理论，植根于统计力学的随机学习机器，最后是与动态规划相关的增强式学习。神经网络动力学模型研究由短期记忆和分层前馈网络构成的动态系统，反馈非线性动态系统的稳定性和联想记忆，以及另一类非线性动态驱动的递归网络系统。

本书注重对数学分析方法和性能优化的讨论，强调神经网络在模式识别、信号处理和控制系统等实际工程问题中的应用。书中包含大量例题和习题，并配有 13 个基于 MATLAB 软件的计算机实验程序。

本书适于作研究生或大学高年级学生的教材，也可作希望深入学习神经网络的科技人员的参考书。

Authorized translation from the English language edition entitled *Neural Networks: A Comprehensive Foundation*, 2nd Edition (ISBN: 0-13-273350-1) by Simon Haykin, published by Pearson Education, Inc., publishing as Prentice Hall PTR, Copyright © 1999 by Prentice-Hall, Inc.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanic, including photocopying, recording, or by any information storage retrieval system, without permission of Pearson Education, Inc.

Chinese simplified language edition published by China Machine Press.

Copyright © 2004 by China Machine Press.

本书中文简体字版由美国 Pearson Education 培生教育出版集团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。

版权所有，侵权必究。

本书版权登记号：图字：01-2001-2207

图书在版编目(CIP)数据

神经网络原理(原书第2版)/海金(Haykin, S.)著；叶世伟等译。-北京：机械工业出版社，2004.1

(计算机科学丛书)

书名原文：Neural Networks: A Comprehensive Foundation, 2nd Edition

ISBN 7-111-12759-5

I. 神… II. ①海… ②叶… III. 人工神经网络 IV. TP183

中国版本图书馆 CIP 数据核字(2003)第 068446 号

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：李伯民 刘 渊

北京牛山世兴印刷厂印刷·新华书店北京发行所发行

2004 年 1 月第 1 版第 1 次印刷

787mm × 1 092mm 1/16 · 41 印张

印数：0 001-4 000 册

定价：69.00 元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换
本社购书热线：(010)68326294

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭橥了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall、Addison-Wesley、McGraw-Hill、Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum、Stroustrup、Kernighan、Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及度藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专诚为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业

的教学度身订造的。其中许多教材均已为M. I. T.、Stanford, U.C. Berkeley, C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程,而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下,读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑,这些因素使我们的图书有了质量的保证,但我们的目标是尽善尽美,而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正,我们的联系方法如下:

电子邮件: hzedu@hzbook.com

联系电话: (010) 68995264

联系地址: 北京市西城区百万庄南街1号

邮政编码: 100037

专家指导委员会

(按姓氏笔画顺序)

尤晋元
石教英
张立昂
邵维忠
周克定
郑国梁
高传善
裘宗燕

王 珊
吕 建
李伟琴
陆丽娜
周傲英
施伯乐
梅 宏
戴 葵

冯博琴
孙玉芳
李师贤
陆鑫达
孟小峰
钟玉琢
程 旭

史忠植
吴世忠
李建中
陈向群
岳丽华
唐世渭
程时端

史美林
吴时霖
杨冬青
周伯生
范 明
袁崇义
谢希仁

译者序

神经网络系统研究的重要意义已为许多科学家所承认，它是智能计算发展的一个主流方向。20 世纪 80 年代中期以来，神经网络重新引起了许多科技工作者的兴趣，形成近代非线性科学和计算智能研究的主要内容之一。特别是神经网络经历了新近 20 年的迅速发展，它所具备的独特知识表示结构和信息处理的原则，使其在许多应用领域取得了显著的进展，能够为解决一些传统计算机极难求解的问题提供满意的解，或者为寻求满意解提供全新的思路。

神经网络由于其信息处理机制和成功应用，实际上已成为智能信息处理的主要技术之一。世界上许多知名大学开设了神经网络的研究生专门课程。在中国，多年以来神经网络也被纳入许多著名大学的研究生课程，使得神经网络这个信息处理工具逐渐为许多智能信息处理工作者所掌握。

在神经网络理论日渐成熟，它的应用逐渐扩大和深入的形势下，如何把握神经网络的研究方向，面向应用和面向广大神经网络的应用者和研究者，介绍神经网络的系统理论和最新发展，成为神经网络课程教学面临的重大挑战。其中如何挑选好的教材成为关键的第一步。目前国内已有的一些神经网络教材，其内容还停留在国际上神经网络上世纪 90 年代初期的发展水平。纵观神经网络发展的历史，特别是近 20 年发展的历史，我们可以发现神经网络的理论和学习算法越来越面向信息处理，它们和生物智能方面建立联系的同时，统计理论、信息理论以及函数空间理论方面的联系日趋紧密。这些方面从最近 10 年的发展看得更清楚。不管是从独立分量分析、支持向量机网络、正则化网络和高斯过程，还是从徐雷教授的阴阳机学习理论，我们都可以发现神经网络的发展目标就是成为智能信息处理的核心工具之一。这一点在模式识别领域已成为现实。作为神经网络的研究生教材应该反映神经网络的这一鲜明特征，但目前国内出版的多数教材尚难达到这一要求。这本书正是反映了神经网络研究的主流发展方向和最新研究内容，所以自出版以来就成为许多国际知名大学的神经网络研究生教材，受到教师和学生们的广泛赞誉。

本书主要讲述神经网络的基本概念，介绍实用的网络模型和学习算法。全书分为 15 章，内容涵盖神经网络理论导论、监督学习、无监督学习和神经网络动力学模型。神经网络导论包括神经元模型和网络结构、机器学习的基本理论。监督学习包括感知机学习规则、有监督的 Hebb 学习、Widrow-Hoff 学习算法、反向传播算法及其变形、RBF 网络、正则化网络、支持向量机网络和委员会机器。无监督学习包括主分量分析、自组织特征映射、用于无监督学习的信息理论、植根于统计力学的随机学习机器和增强式学习。神经动力学模型包括嵌入短时记忆的分层前馈网络动力系统、递归网络的稳定性和学习问题及其在联想记忆中的应用。书中注重对数学分析方法和性能优化的讨论，强调神经网络在模式识别、信号处理以及控制系统等实际工程问题中的应用。同时本书包含大量例题和习题，并配有 13 个基于 MATLAB 软件的计算机实验的源程序。

清华大学出版社已经出版本书的影印版，这使得中国的学生有幸直接阅读英文原版教

材。但是译者在中国科学院研究生院使用本书影印版作为教材讲授神经网络的过程中发现,一方面由于研究生课时的限制,在一学期内全部讲授本书内容极其困难,所以只能选择其中一些内容讲授;另一方面,由于本书材料非常丰富及其完整性,部分讲解书中内容对于系统学习和掌握神经网络的原理和研究方法是不利的,而初学者要完整阅读本书原文困难也很大。另外,由于神经网络的应用逐渐深入和扩大,许多学习神经网络的其他读者其实仅仅需要了解神经网络的基本原理和系统方法,他们学习神经网络的目的是为了从中找到具体应用领域的解决方法或者获得解决问题的新思路。对于他们而言,直接阅读原著是不现实的,而且没有必要。基于上述原因,翻译出版本书是很有必要的。

由于神经网络的迅速发展,许多神经网络的新名词和概念还没有确定的中文翻译,所以在本书中凡是我们认为不能完全确定的名词或术语都在其第一次出现的地方给出对应的英语词汇,有一些地方甚至直接引用英语词汇本身。最后在书后还有中英文对照索引。

在这本书的翻译中,我们力求忠实、准确地反映原著的内容,同时也力求保留原著的风格。但由于神经网络属于多学科交叉领域,研究范围很广,近年来研究成果层出不穷,而且译者水平有限,书中错误和不准确之处在所难免,恳请读者批评指正。

致 谢

本书的翻译得到了中国科学院研究生院院长基金的资助(项目编号 YZJJ200206)。

在这里,我们要感谢机械工业出版社华章公司组织翻译出版这本重要著作,感谢编辑们的细心加工和修改,没有这些本书的出版是不可想像的。感谢中国科学院研究生院 2001 级和 2002 级选修神经网络课程的研究生,正是他们的需求才促使我们翻译这本教材。同时我们还要感谢中国科学院计算技术研究所智能信息处理重点实验室的支持。

译 者

2003 年 5 月于北京

前 言

神经网络,或者更精确地说人工神经网络,是一种植根于许多学科的技术,其中涉及神经科学、数学、统计学、物理学、计算机科学和工程学。神经网络具有的一个重要性质,即在有教师或无教师的情况下能够从输入数据中进行学习的能力,这使得它在不同领域中得到应用,如建模、时间序列分析、模式识别、信号处理和控制。

由于神经网络具有多学科性,本书对该主题进行了综合论述,并给出了大量例子、基于计算机的实验、习题以及参考文献进行补充分析。

本书由四部分组成,组织如下:

1. 介绍材料,由第1章和第2章组成。第1章大体上定性描述什么是神经网络,它们的性质、组成及其怎样和人工智能相联系。这章以一些历史注释结束。第2章提供学习过程的许多侧面的概述及其统计性质。该章引进了一个重要概念,即 Vapnik-Chervonenkis (VC) 维数,用于度量学习机器所实现的一簇分类函数的容量。

2. 有教师学习机器,由第3章至第7章组成。第3章研究这部分中最简单的神经网络:涉及一个或多个输出神经元但无隐藏神经元的网络。该章描述最小均方(LMS)算法(在设计线性自适应滤波器时非常流行)和感知器收敛定理。第4章给出利用反向传播算法训练的多层感知器的完全处理。这个算法(代表 LMS 算法的一种推广)已经作为神经网络的推进器而出现。第5章给出另一类分层神经网络即径向基函数网络详细的数学处理,它们的构成包括一层基函数。这一章强调在设计 RBF 网络中正则化理论的作用。第6章描述一类比较新的学习机器,即支持向量机,它的理论建立在第2章给出的统计学习理论的材料上。本书第二部分以第7章结束,讨论委员会机器,它的构成包括几个学习者作为组成部分。在这一章我们描述总体平均,推举和分层混合专家三种不同的构建委员会机器的方法。

3. 无教师学习机器,由第8章至第12章组成。第8章把 Hebb 学习应用到主分量分析。第9章把自组织学习的另一形式,也就是竞争学习,应用于构造以自组织映射著称的计算映射。这两章突出强调学习规则根植于神经生物学。第10章注意于设计无监督学习算法的信息理论,强调它们在建模、图像处理和独立分量分析中的应用。第11章描述植根于和信息理论有密切关系的统计力学的自监督学习机器。第12章,介绍动态规划和它与增强式学习的关系。

4. 非线性动态系统,由第13章至15章组成。第13章描述一类由短期记忆和分层前馈网络结构组成的动态系统。第14章强调涉及使用反馈的非线性动态系统所引起的稳定性问题。该章还讨论联想记忆的例子。第15章描述另一类非线性动态系统,即递归网络,它依赖于使用反馈完成输入-输出映射。

本书后记简要描述神经网络在构造用于模式识别、控制和信号处理的智能机器时所起的作用。

本书的组织在神经网络研究生课程的使用上给予了很大灵活性,教师可根据需要灵活选择讲课内容。全书中总共包括15个基于计算机的实验,其中有13个实验需使用 MATLAB。

MATLAB 实验的文件可直接从以下网站下载:

<ftp://ftp.mathworks.com/pub/books/haykin>

<http://www.mathworks.com/books/>

每章后都附有习题。许多习题具有挑战性, 不仅能检查本书的使用者对本书所包含的资料掌握的程度, 而且扩充了这些资料。

工程师、计算机科学家和物理学家也会从本书获益。希望本书对其他学科, 如心理学和神经科学的研究人员, 也会有所帮助。

Simon Haykin

于 Hamilton, Ontario

1998 年 2 月

缩写和符号

缩写

AI	artificial intelligence 人工智能
APEX	adaptive principal components extraction 自适应主分量分析
AR	autoregressive 自回归
BBTT	back propagation through time 通过时间的反向传播
BM	Boltzmann machine Boltzmann 机
BP	back propagation 反向传播
b/s	bits per second 每秒比特率
BOSS	bounded, one-sided saturation 有界, 单边饱和
BSB	brain-state-in-a-box 盒中脑状态
BSS	Blind source (signal) separation 盲源(信号)分离
CART	classification and regression tree 分类和回归树
cmm	correlation matrix memory 相关矩阵记忆
CV	cross-validation 交叉确认
DEKF	decoupled extended Kalman filter 解耦扩展 Kalman 滤波器
DFA	deterministic finite-state automata 确定性有限状态自动机
DSP	digital signal processor 数字信号处理器
EKF	extended Kalman filter 扩展 Kalman 滤波器
EM	expectation-maximization 期望最大化
FIR	finite-duration impulse response 有限时间冲击响应
FM	frequency-modulated (signal) 频率调制(信号)
GEKF	global extended Kalman filter 全局扩展 Kalman 滤波器
GCV	generalized cross-validation 广义交叉确认
GHA	generalized Hebbian algorithm 广义 Hebb 算法
GSLC	generalized sidelobe canceler 广义旁瓣消除器
HME	hierarchical mixture of expert 分层混合专家

HMM	hidden Markov model 隐 Markov 模型
Hz	hertz 赫兹
ICA	independent component analysis 独立分量分析
Infomax	maximum mutual information 最大互信息
KR	kernel regression 核回归
LMS	least-mean-square 最小均方
LR	likelihood ratio 似然比
LTP	long-term potentiation 长期电位(LPT)
LTD	long-term depression 长期衰减
LR	likelihood ratio 似然比
LVQ	learning vector quantization 学习向量量化
MCA	minor component analysis 次分量分析
MDL	minimum description length 最小描述长度
ME	mixture of expert 混合专家
MFT	mean-field theory 平均场理论
MIMO	multiple input-multiple output 多输入多输出
ML	maximum likelihood 最大似然
MLP	multilayer perceptron 多层感知器
MRAC	model reference adaptive control 模型参考自适应控制
NARMA	nonlinear autoregressive moving average 非线性自回归滑动平均
NARX	nonlinear autoregressive with exogenous input 具有外部输入的非线性自回归
NDP	neuron-dynamic programming 神经动态规划
NW	Nadaraya-Watson (estimator) Nadaraya-Watson(估计器)
NWKR	Nadaraya-Watson kernel regression Nadaraya-Watson 核回归
OBD	optimal brain damage 最优脑损伤
OBS	optimal brain surgeon 最优脑外科
OCR	optical character recognition 光学字符识别
ODE	ordinary differential equation 常微分方程
PAC	probably approximately correct 可能近似正确
PCA	principal component analysis 主分量分析
pdf	probability density function 概率密度函数
pmf	probability mass function 概率质量函数

RBF	radial basis function 径向基函数
RMLP	recurrent multilayer perceptron 递归多层感知器
RTRL	real-time recurrent learning 实时递归学习
SIMO	single input-multiple output 单输入多输出
SISO	single input-single output 单输入单输出
SNR	signal-to-noise ratio 信噪比
SOM	self-organizing map 自组织映射
SRN	simple recurrent network(also referred to as Elman's recurrent network) 简单递归网络(也称为 Elman 递归网络)
SVD	singular value decomposition 奇异值分解
SVM	support vector machine 支持向量机
TDNN	time-delay neural network 时延神经网络
TLFN	time lagged feedforward network 时间滞后前馈网络
VC	Vapnik-Chervononkis (dimension) Vapnik-Chervononkis(维数)
VLSI	very-large-scale integration 超大规模集成
XOR	exclusive OR 异或

重要的符号

a	action 动作
$\mathbf{a}^T \mathbf{b}$	inner product of vectors $\mathbf{a}$ and $\mathbf{b}$ 向量 $\mathbf{a}$ 和 $\mathbf{b}$ 的内积
$\mathbf{a} \mathbf{b}^T$	output product of vectors $\mathbf{a}$ and $\mathbf{b}$ 向量 $\mathbf{a}$ 和 $\mathbf{b}$ 的外积
$\binom{l}{m}$	binomial coefficient 二项式系数
$A \cup B$	unions of A and B A 和 B 的并
B	inverse of temperature 温度的逆
b_k	bias applied to neuron k 神经元 k 的偏置
$\cos(\mathbf{a}, \mathbf{b})$	cosine of the angle between vectors $\mathbf{a}$ and $\mathbf{b}$ 向量 $\mathbf{a}$ 和 $\mathbf{b}$ 夹角的余弦
D	depth of memory 记忆深度
$D_{f\parallel g}$	Kullback-Leibler divergence between probability density functions f and g 概率密度函数 f 和 g 之间的 Kullback-Leibler 散度
$\tilde{\mathbf{D}}$	adjoint of operator $\mathbf{D}$ 算子 $\mathbf{D}$ 的伴随
E	energy function 能量函数
E_i	energy of state i in statistical mechanics 统计力学中状态 i 的能量
E	statistical expectation operator 统计期望算子

$\langle E \rangle$	average energy 平均能量
erf	error function 误差函数
erfc	complimentary error function 误差函数的补
exp	exponential 指数
$\mathcal{E}_{av}$	average squared error or sum of squared error 平均平方误差或平方误差和
$\mathcal{E}(n)$	instantaneous value of the sum of squared error 平方误差和的瞬时值
$\mathcal{E}_{total}$	total sum of error squares 总平方误差和
F	free energy 自由能量
$f_X(\mathbf{X})$	probability density function of random vector $\mathbf{X}$ 随机向量 $\mathbf{X}$ 的概率密度函数
$\mathcal{F}^*$	subset (network) with the smallest minimum empirical risk 经验风险最小值最小的子集(网络)
$\mathbf{H}$	Hessian matrix Hessian 矩阵
$\mathbf{H}^{-1}$	inverse of matrix $\mathbf{H}$ 矩阵 $\mathbf{H}$ 的逆
i	square root of -1 , also denoted by j -1 的平方根, 亦记作 j
$\mathbf{I}$	identity matrix 单位矩阵
$\mathbf{I}$	Fisher's information matrix Fisher 信息矩阵
J	mean-square error 平均平方误差
$\mathbf{J}$	Jacobian matrix Jacobi 矩阵
$\mathbf{K}(n, n-1)$	error covariance matrix in Kalman filter theory Kalman 滤波理论中的误差协方差矩阵
$\mathbf{K}^{1/2}$	square root of matrix $\mathbf{K}$ 矩阵 $\mathbf{K}$ 的平方根
$\mathbf{K}^{T/2}$	transpose of square root of matrix $\mathbf{K}$ 矩阵 $\mathbf{K}$ 的平方根的转置
k_B	Boltzmann constant Boltzmann 常数
log	logarithm 对数
$L(\mathbf{w})$	log-likelihood function of weight vector $\mathbf{w}$ 权值向量 $\mathbf{w}$ 的对数似然函数
$\mathcal{L}(\mathbf{w})$	log-likelihood function of weight vector $\mathbf{w}$ based on a single example 单样本的权值向量 $\mathbf{w}$ 的对数似然函数
$\mathbf{M}_c$	controllability matrix 可控性矩阵
$\mathbf{M}_o$	observability matrix 可观察性矩阵
n	discrete time 离散时间
p_i	probability of state i in statistical mechanics 统计力学中状态 i 的概率
p_{ij}	transition probability from state i to state j 从状态 i 到状态 j 的转移概率
$\mathbf{P}$	stochastic matrix 随机矩阵
P_c	probability of correct classification 正确分类的概率
P_e	probability of error 误差概率
$P(e \mathcal{C})$	conditional probability of error e given that the input is drawn from class $\mathcal{C}$ 从类 $\mathcal{C}$ 中输入时误差 e 的条件概率
p_α^+	probability that the visible neurons of a Boltzmann machine are in state α , given that the

	network is in its clamped condition(i.e., positive phase) 假设网络处于箝制条件(即正向阶段)时, Boltzmann 机的可见神经元状态为 α 的概率
p_{α}^{-}	probability that the visible neurons of a Boltzmann machine are in state α , given that the network is in its free-running condition(i.e., negative phase) 假设网络处于自由运行条件(即负向阶段)时, Boltzmann 机的可见神经元状态为 α 的概率
$\hat{r}_x(j, k; n)$	estimate of autocorrelation function of $x_j(n)$ and $x_k(n)$ $x_j(n)$ 和 $x_k(n)$ 的自相关函数估计
$\hat{r}_{dx}(k; n)$	estimate of cross-correlation function of $d(n)$ and $x_k(n)$ $d(n)$ 和 $x_k(n)$ 的交叉相关函数估计
R	correlation matrix of an input vector 输入向量的相关矩阵
t	continuous time 连续时间
T	temperature 温度
$\mathcal{T}$	training set(sample) 训练集(样本)
tr	trace of a matrix operator 矩阵算子的迹
var	variance operator 方差算子
$V(\mathbf{x})$	Lyapunov function of state vector $\mathbf{x}$ 状态向量 $\mathbf{x}$ 的 Lyapunov 函数
v_j	induced local field or activation potential of neuron j 神经元 j 的诱导局部域或激活位势
$\mathbf{w}_o$	optimum value of synaptic weight vector 突触权值向量的最优值
w_{kj}	synaptic weight of synapse j belonging to neuron k 属于神经元 k 的突触 j 的突触权值
$\mathbf{w}^*$	optimum weight vector 最优权植向量
$\bar{\mathbf{x}}$	equilibrium value of state vector $\mathbf{x}$ 状态向量 $\mathbf{x}$ 的平衡值
$\langle x_j \rangle$	average of state x_j in a "thermal" sense "热"意义下状态 x_j 的平均
$\hat{x}$	estimate of x , signified by the use of a caret(hat) x 的估计, 用加字符号^(帽符号)表示
$ x $	absolute value(magnitude) of x x 的绝对值(幅度)
x^*	complex conjugate of x , signified by asterisk as superscript 状态 x 的复共轭, 用星号 * 作上标
$\ \mathbf{x}\ $	Euclidean norm (length) of vector $\mathbf{x}$ 向量 $\mathbf{x}$ 的欧几里德范数(长度)
$\mathbf{x}^T$	transpose of vector $\mathbf{x}$, signified by the superscript T 向量 $\mathbf{x}$ 的转置, 用上标 T 表示
z^{-1}	unit delay operator 单位延迟算子
Z	partition function 剖分函数
$\delta_j(n)$	local gradient of neuron j at time n 神经元 j 在时刻 n 的局部梯度
Δw	small change applied to weight w 权值 w 的微小改变
∇	gradient operator 梯度算子
∇^2	Laplacian operator Laplace 算子
$\nabla_w J$	gradient of J with respect to w J 关于 w 的梯度

$\nabla \cdot \mathbf{F}$	divergence of vector $\mathbf{F}$	向量 $\mathbf{F}$ 的散度
η	learning-rate parameter	学习率参数
κ	cumulant	累积量
μ	policy	策略
θ_k	threshold applied to neuron k (i.e., negative of bias b_k)	神经元 k 的國值(即偏置 b_k 的负值)
λ	regularization parameter	正则化参数
λ_k	k th eigenvalue of a square matrix	方阵的第 k 个特征值
$\varphi_k(\cdot)$	nonlinear activation function of neuron k	神经元 k 的非线性激活函数
$\in$	symbol for “belong to”	“属于”符号
$\cup$	symbol for “union of”	“并”符号
$\cap$	symbol for “intersection of”	“交”符号
$*$	symbol for convolution	“卷积”符号
$+$	superscript symbol for pseudoinverse of a matrix	矩阵伪逆的上标符号

开区间和闭区间

- 变量 x 的开区间 (a, b) 表示 $a < x < b$ 。
- 变量 x 的闭区间 $[a, b]$ 表示 $a \leq x \leq b$ 。
- 变量 x 的半闭半开区间 $[a, b)$ 表示 $a \leq x < b$ ；类似地，变量 x 的半开半闭区间 $(a, b]$ 表示 $a < x \leq b$ 。

最小和最大

- 符号 $\arg \min_{\mathbf{w}} f(\mathbf{w})$ 表示函数 $f(\mathbf{w})$ 关于变元向量 $\mathbf{w}$ 的最小值。
- 符号 $\arg \max_{\mathbf{w}} f(\mathbf{w})$ 表示函数 $f(\mathbf{w})$ 关于变元向量 $\mathbf{w}$ 的最大值。

目 录

出版者的话		第3章 单层感知器	81
专家指导委员会		3.1 简介	81
译者序		3.2 自适应滤波问题	82
前言		3.3 无约束最优化技术	83
缩写和符号		3.4 线性最小二乘滤波器	87
第1章 导言	1	3.5 最小均方算法	89
1.1 什么是神经网络	1	3.6 学习曲线	92
1.2 人脑	4	3.7 学习率退火进度	93
1.3 神经元模型	7	3.8 感知器	94
1.4 看作有向图的神经网络	10	3.9 感知器收敛定理	94
1.5 反馈	12	3.10 Gauss 环境下感知器与 Bayes 分类器 的关系	99
1.6 网络结构	13	3.11 小结和讨论	102
1.7 知识表示	15	注释和参考文献	103
1.8 人工智能和神经网络	22	习题	105
1.9 历史注释	24	第4章 多层感知器	109
注释和参考文献	29	4.1 简介	109
习题	30	4.2 预备知识	111
第2章 学习过程	33	4.3 反向传播算法	112
2.1 简介	33	4.4 反向传播算法小结	121
2.2 误差修正学习	34	4.5 异或问题	123
2.3 基于记忆的学习	35	4.6 改善反向传播算法性能的试探法	125
2.4 Hebb 学习	36	4.7 输出表示和决策规则	129
2.5 竞争学习	39	4.8 计算机实验	131
2.6 Boltzmann 学习	40	4.9 特征检测	140
2.7 信任赋值问题	41	4.10 反向传播和微分	142
2.8 有教师学习	42	4.11 Hessian 矩阵	143
2.9 无教师学习	43	4.12 泛化	144
2.10 学习任务	44	4.13 函数逼近	146
2.11 记忆	50	4.14 交叉确认	150
2.12 自适应	56	4.15 网络修剪技术	154
2.13 学习过程的统计性质	57	4.16 反向传播学习的优点和局限	160
2.14 统计学习理论	60	4.17 反向传播学习的加速收敛	165
2.15 可能近似正确的学习模型	69	4.18 作为最优化问题看待的有监督 学习	165
2.16 小结和讨论	72	4.19 卷积网络	173
注释和参考文献	73	4.20 小结和讨论	175
习题	77		