

华南理工大学科学丛书

神经网络 理论与应用

徐秉铮 张百灵 韦岗 编著

华南理工大学出版社

华南理工大学科学丛书

神经网络理论与应用

徐秉铮 张百灵 韦 岗 编著

华南理工大学出版社

图书在版编目 (CIP) 数据

神经网络理论与应用/徐秉铮等编著. —广州: 华南理工大学出版社, 1994. 12

(华南理工大学科学丛书)

ISBN 7-5623-0686-9

I. 神…

Ⅱ. 徐…

Ⅲ. ①神经网络-仿真网络 ②神经网络-计算机网络

Ⅳ. ①TP18 ②TP393

华南理工大学出版社出版发行

(广州五山·邮编 510641)

责任编辑 杨昭茂

各地新华书店经销

华南理工大学出版社电脑室排版

广东封开人民印刷厂印装

1994年12月第1版 1994年12月第1次印刷

开本: 850×1168 1/32 印张: 17.5 字数: 439千

印数: (精) 1—500 (半精) 1—2 500

定价: (精) 22.00元 (半精) 18.00元

前 言

人工神经网络或神经计算是近十几年来发展极为迅速的一门边缘学科。它在信号处理、模式识别、智能控制等领域中越来越多的成功应用是令人鼓舞的，即使是最初持观望态度的人也不得不逐渐放弃对它的怀疑。神经网络对众多学科的包容性、应用范围的广泛性以及理论分析方法的多样性是前所未有的。Hopfield 将能量函数引入对称、反馈的网络结构，研究联想记忆和最优化计算，以及 Rumelhart, Hinton 和 Williams 提出反向传播算法训练多层前传网络以逼近期望的连续函数，是神经网络发展历史中永远值得记载的重大事件。正是这两项研究重新唤起了人们对神经网络的研究热情，促进了神经网络的迅猛发展。虽然自此以后很难评价神经网络的哪一项研究标志着新的重大突破，但我们可以越来越清楚地看到，神经网络理论的研究正越来越深化，其框架体系日臻完善；神经网络硬件的快速发展以及在各个领域中应用的成功经验，正促使它向实用化转变；同时，神经网络软件和专用处理器的出现，也使得神经计算机不再是神话。当然，这并不是说神经网络已经发展成熟，相反，新的网络结构、新的学习算法、新的理论研究手段仍在层出不穷。科学史上很少出现过这样一门学科能带来如此众多的新思维、新概念。其实，神经网络的这一特点是不足为奇的，因为它的发展的原动力在于对人类大脑和智能的研究，而在这方面，也许我们至今只能算看见了新大陆的海岸线，已经认识的和尚未认识的相比是微不足道的。尽管我们很难预测神经网络的发展前景，但可以肯定的是，它绝不会再重遭 60 年代被“打倒”的命运，道路也许很曲折，但必定产生

mxas/04

一个又一个的飞跃。

本书是为信息处理类专业的研究生编写的，因此我们假定读者已熟练掌握矩阵运算、微积分和概率统计等数学知识。全书除第一章绪论外，主要包括三部分：第一部分第二章至第四章，主要以 Hopfield 模型为线索，介绍神经网络实现联想记忆、模式识别和最优化计算的基本模型和方法。这一部分侧重模型的结构和基本原理，很少涉及学习或训练问题（这里我们将编码与学习或训练区别开来）。第二部分第五章至第八章，围绕多层感知器，讨论有监督学习算法和相应的理论问题，并适当介绍一些简单的应用实例。第三部分第九章和第十章是关于无监督学习问题，其中第十章的无监督竞争学习由于其内容极为广泛（如 ART、认知机等），我们只能作一些简单介绍。值得指出的是，第一，无监督学习和有监督学习之间并没有明确的分界线，在一个模型或某个应用中完全可以同时利用两类算法；第二，一种学习算法的提出常常是针对一种特殊的模型结构，但并不一定局限于该结构。如果读者在学习一种算法时将其推广，发现问题，解决问题，那么对神经网络的深入认识是很有帮助的；第三，限于篇幅，我们在本书中没有介绍模糊神经网络。

本书的编写主要参考了 J. Hertz, A. Krogh 和 R. G. Palmer 编著的“Introduction to the Theory of Neural Computation”一书以及近年来发表在一些国际专业杂志（如 Neural Computation, Neural Networks, IEEE Transactions on Neural Networks 等）和重要的国际会议上的一些学术论文，也结合了作者近年来的一些研究成果。神经网络研究的学术论文每年数以千计，要想在我们这样一本书中面面俱到地介绍神经网络理论和应用的各个方面，几乎是不可能的事，我们只能尽自己的热情和努力以及手头所掌握的材料，在阐述神经网络原理的同时，尽可能全面地反映出神经网络的概貌和体系，尽可能多地推荐给读者一些我们认为

值得重视的新课题和研究动向。由于客观条件的限制，出版时间的紧迫，以及作者水平和能力的局限，本书与作者的愿望仍有很大差距。因此，我们期待着通过教学，通过读者的热情批评，使书中差错能够得到纠正，在可能的再版中内容得到充实和完善。

本书的出版得到华南理工大学专项基金的资助。第二作者感激家人对他在为本书所付出的劳动过程中所给予的关怀、理解和帮助。

编 著 者
1994 年 3 月

《华南理工大学科学丛书》编审委员会

主 任:刘振群

副主任:刘正义

委 员 (以姓氏笔划为序)

刘正义 刘有然 刘振群

刘焕彬 江厚祥 张力田

沈尧天 周泽华 周绍华

林维明 徐秉铮 贾信真

韩大健

出版说明

华南理工大学,是国家教育委员会直属的一所全国重点高等学校。自1952年建校以来,尤其是改革开放十多年来,已发展成为以工为主、理工文管结合的颇具规模的多科性综合大学,为国家培养和输送了6万多名各学科、专业的高级建设人才。目前,在校任教的教授、副教授(含相当的其它高级职称人员)有1 000多人。正是他们带领4 500多名教职员工,沿着党的教育方针指引的道路,把学校办成既是教学中心,又是科研中心。在他们当中有相当一批是在基础理论、应用科学或工程技术领域方面探索研究几十年而取得可喜成果,为繁荣发展我国科学和教育事业作出重要贡献并在国内外享有声誉的著名教授、专家、学者。我们编辑出版《华南理工大学科学丛书》的宗旨是:选择华南理工大学任职的教授、专家、学者科学研究中在某学科或领域处于科学前沿或取得突破性成果的科学论著,按统一的规格要求汇编成丛书,陆续出版面世,以集中反映华南理工大学代表性的科研成果及其在有关学科领域中的地位 and 特色,为科研成果的积累、传播、交流及转化为生产方面提供必要的条件。

《华南理工大学科学丛书》列选资格由本丛书编审

委员会审定,并由华南理工大学拨出专项基金资助出版。选编的原则是:

一、列入本丛书的作者(或第一作者),必须是在华南理工大学担任现职的教师或在学的研究生。除博士点或博士后流动站的导师外,其他教师或研究生的著作必须有两名以上同行专家提出评论意见推荐。

二、凡申请列入本丛书的著作,均应是反映华南理工大学某学科领域的优势、特色和科学水平的,在国内或国内外处于先进水平或领先地位的学术专著;对科学研究成果卓著、学术界知名度大且年事已高的老教授,有选择地适当安排出版其科学论文集。

三、列入本丛书的著作均应使用中华人民共和国法定计量单位(或国际单位制)和有关的新标准(1990年前发表的著作附新制与旧制对照表);在引用他人著作或观点时,应注明出处和列出参考文献。

四、作者已在国外发表的论著,选用时保留发表时的文种,并注明发表的时间、地点和刊物名称。

出版高层次的科学丛书,我们还是初次尝试,由于缺乏经验,如出现缺点和错误,欢迎读者批评指正。

华南理工大学出版社

目 录

前言

第一章 绪论	1
§ 1. 1 概述	1
§ 1. 1. 1 神经元	3
§ 1. 1. 2 并行处理	7
§ 1. 2 神经网络的研究简史	9
§ 1. 3 神经网络的概念、定义与基本组成	15
§ 1. 4 神经网络的研究范畴	19
§ 1. 5 神经网络模型的一些研究方法	26
第二章 联想记忆的基本原理	33
§ 2. 1 线性联想映射	35
§ 2. 1. 1 Kohonen 模型	36
§ 2. 1. 2 利用最佳秩 r -逼近的线性联想记忆	41
§ 2. 1. 3 最小平方联想记忆	43
§ 2. 2 Hopfield 联想记忆模型	44
§ 2. 2. 1 Hopfield 模型连接权的确定	45
§ 2. 2. 2 Hopfield 模型的存储容量	50
§ 2. 2. 3 Hopfield 联想记忆模型的匹配滤波器实现	55
§ 2. 3 Hopfield 模型中的一些稳定性问题	60
§ 2. 3. 1 稳定性和吸引性的概念	60
§ 2. 3. 2 能量函数	64
§ 2. 3. 3 保证稳定性的几个充分条件	67
§ 2. 4 Hopfield 模型的一些变形与推广	70
§ 2. 4. 1 Hopfield 模型的几种变形	71
§ 2. 4. 2 非正交模式	76

§ 2. 4. 3 连续值单元	80
§ 2. 4. 4 模式的时态序列	87
第三章 联想记忆与模式识别的实用模型	93
§ 3. 1 引入三态隐单元的 Hopfield 模型	93
§ 3. 1. 1 能量函数峰值的正交性	95
§ 3. 1. 2 样本存贮过程	96
§ 3. 1. 3 样本恢复过程	98
§ 3. 2 Kanerva 稀疏分布式联想记忆	100
§ 3. 3 指数响应单元网络 (ECAM)	106
§ 3. 4 分形神经网络 (FNN)	113
§ 3. 4. 1 网络结构	114
§ 3. 4. 2 信息的存贮和检索	114
§ 3. 5 交替投影神经网络 (APNN)	118
§ 3. 6 元胞神经网络 (CNN)	123
§ 3. 7 双向联想记忆 (BAM) 模型	130
§ 3. 7. 1 BAM 的结构	131
§ 3. 7. 2 简单 BAM 模型的性能分析	135
§ 3. 7. 3 BAM 的应用实例	139
§ 3. 8 BAM 模型的改进	143
§ 3. 8. 1 BAM 的多重训练算法	143
§ 3. 8. 2 具有指数容量的双向联想记忆	145
第四章 优化计算的神经网络方法	150
§ 4. 1 加权匹配问题	152
§ 4. 2 旅行售货员 (TSP) 问题	156
§ 4. 2. 1 Hopfield-Tank 方法	157
§ 4. 2. 2 Hopfield-Tank 方法的一种改进	159
§ 4. 2. 3 求解 TSP 的弹性网络方法	161
§ 4. 2. 4 N-皇后问题	163
§ 4. 3 图划分问题	166
§ 4. 3. 1 图二分问题	167

§ 4. 3. 2 图划分问题	169
§ 4. 4 动态规划 (DPP) 问题	170
§ 4. 4. 1 最优性原理	170
§ 4. 4. 2 动态规划的人工神经网络	172
§ 4. 4. 3 网络仿真	175
§ 4. 5 求解优化问题的马尔可夫神经网络	177
§ 4. 6 求解一般非线性优化问题的广义 Hopfield 网络	183
§ 4. 6. 1 无约束优化问题	183
§ 4. 6. 2 增广拉格朗日方法	186
§ 4. 6. 3 广义缩减梯度法	187
§ 4. 6. 4 逐次二次规划 (SQP)	188
§ 4. 7 优化计算神经网络的应用实例	191
§ 4. 7. 1 基于优化计算神经网络的模/数 (A/D) 转换器	191
§ 4. 7. 2 图象处理中的优化问题	195
第五章 感知器与 Madaline 的基本原理	204
§ 5. 1 基本概念	205
§ 5. 1. 1 自适应线性组合器	207
§ 5. 1. 2 线性分类器——单阈值元件	208
§ 5. 1. 3 线性可分性	210
§ 5. 1. 4 非线性分类器	213
§ 5. 2 误差校正规则	217
§ 5. 2. 1 线性规则	218
§ 5. 2. 2 非线性规则	219
§ 5. 2. 3 感知器学习规则的几何解释	222
§ 5. 2. 4 感知器学习规则的收敛性证明	225
§ 5. 2. 5 Madaline 网络的误差校正规则	227
§ 5. 3 线性单元与非线性单元	230
§ 5. 3. 1 线性单元	230
§ 5. 3. 2 非线性单元	236
§ 5. 3. 3 随机单元	238
§ 5. 4 简单感知器的容量	240

§ 5. 5 Madaline 网络中的最陡下降学习	244
第六章 反向传播学习算法及其应用	249
§ 6. 1 反向传播算法	250
§ 6. 2 反向传播算法的修正与推广	256
§ 6. 3 多层前传网络的性能	265
§ 6. 3. 1 多层前传网络的泛函能力与隐单元的必要个数	266
§ 6. 3. 2 输入表示问题	270
§ 6. 3. 3 推广 (generalization)	272
§ 6. 3. 4 学习的复杂性问题	274
§ 6. 4 关于推广的理论框架	276
§ 6. 4. 1 平均推广能力	277
§ 6. 4. 2 对差的推广概率的定界	283
§ 6. 5 网络结构的优化	288
§ 6. 5. 1 修剪与权的衰减	289
§ 6. 5. 2 网络构造算法	292
§ 6. 6 例子与应用	297
第七章 径向基函数网络	304
§ 7. 1 径向基函数网络的基本原理	305
§ 7. 1. 1 用于插值的径向基函数	305
§ 7. 1. 2 用于模式识别的径向基函数	305
§ 7. 1. 3 RBF 网络	309
§ 7. 2 径向基函数网络的 k-均值聚类算法	312
§ 7. 3 径向基函数网络的正交最小平方学习算法	316
§ 7. 3. 1 RBF 网络基函数中心选取的 OLS 方法	317
§ 7. 3. 2 RBF 网络在通信信道均衡中的应用	322
§ 7. 4 径向基函数插值的资源分配网络	325
§ 7. 5 高斯条 (Gaussian bar) 函数网络	329
§ 7. 6 径向基函数网络的推广	332
§ 7. 6. 1 关于 RBF 网络推广能力的讨论	332
§ 7. 6. 2 预测网络推广性质的基本方法	333

§ 7. 6. 3 改善 RBF 网络推广性质的一种方法	334
§ 7. 7 与径向基函数类似的其它一些模式分类方法	337
§ 7. 7. 1 修正的 Kanerva 模型	337
§ 7. 7. 2 位函数法	339
§ 7. 7. 3 核判别式分析	339
§ 7. 7. 4 球面分级神经网络	340
§ 7. 7. 5 概率神经网络	341
§ 7. 8 径向基函数插值的正则化理论与正则化网络	344
§ 7. 8. 1 正则化理论的基本概念	344
§ 7. 8. 2 正则化网络	346
§ 7. 8. 3 正则化方法的推广——超基函数网络	348
§ 7. 9 其他模式分类器简介	350
第八章 递归神经网络	354
§ 8. 1 递归网络与递归反向传播	355
§ 8. 1. 1 递归网络的基本结构	355
§ 8. 1. 2 连续时间递归网络的递归反向传播算法	358
§ 8. 1. 3 连续时间递归网络的时变递归反向传播算法	364
§ 8. 2 学习时间序列	367
§ 8. 2. 1 时延神经网络 (TDNN)	368
§ 8. 2. 2 部分递归网络	371
§ 8. 2. 3 随时间演化的反向传播 (BPTT)	374
§ 8. 3 递归网络的几种实时学习算法	376
§ 8. 3. 1 实时递归学习 (RTRL)	376
§ 8. 3. 2 随时间演化反向传播的实时算法	379
§ 8. 3. 3 训练递归网络的广义卡尔曼滤波算法	382
§ 8. 4 玻尔兹曼机	386
§ 8. 4. 1 随机单元	386
§ 8. 4. 2 Boltzmann 机的一些推广	391
§ 8. 4. 3 确定性 Boltzmann 机	393
§ 8. 5 玻尔兹曼机的信息几何理论简介	394
§ 8. 5. 1 所有概率分布的流形	394

§ 8. 5. 2 无隐单元的 Boltzmann 机	400
§ 8. 6 强化学习	403
§ 8. 6. 1 联想赏-罚的基本概念	405
§ 8. 6. 2 联想赏-罚理论	408
§ 8. 6. 3 建模网络与评判网络	411
第九章 无监督 Hebb 学习	415
§ 9. 1 概述	415
§ 9. 1. 1 有监督学习和无监督学习的区别	415
§ 9. 1. 2 无监督学习的意义	417
§ 9. 2 无监督 Hebb 学习的基本原理	420
§ 9. 2. 1 关于线性神经元的一些结论	420
§ 9. 2. 2 无监督 Hebb 学习的 Oja 规则	423
§ 9. 2. 3 Oja 规则的理论分析	425
§ 9. 3 主分量分析 (PCA) 学习算法	428
§ 9. 3. 1 PCA 的基本原理	428
§ 9. 3. 2 主分量分析的单层前传网络	432
§ 9. 3. 3 PCA 学习的其它网络结构	439
§ 9. 4 PCA 学习的推广	443
§ 9. 4. 1 自适应主分量抽取算法	443
§ 9. 4. 2 次分量抽取与最优拟合	448
§ 9. 4. 3 其它与 PCA 有关的问题	457
§ 9. 5 模式识别的子空间学习算法简介	458
§ 9. 5. 1 子空间模式识别原理	458
§ 9. 5. 2 双子空间模式识别	462
§ 9. 6 自组织特征抽取	464
§ 9. 6. 1 自组织多层感知器	464
§ 9. 6. 2 仿真结果	468
§ 9. 6. 3 最大信息传输问题	470
第十章 无监督竞争学习	473
§ 10. 1 简单的竞争学习	474

§ 10. 1. 1 赢者取全 (WTA) 的基本原理	474
§ 10. 1. 2 价值函数与收敛性	478
§ 10. 2 竞争学习的例子与应用	480
§ 10. 3 自适应共振理论	487
§ 10. 3. 1 稳定性-可塑性二难问题与 ART 的提出	488
§ 10. 3. 2 ART 的基本原理	489
§ 10. 3. 3 ART 用作分类器时的算法	496
§ 10. 3. 4 ART 的一种简单网络实现	498
§ 10. 4 特征映射的理论和应用	500
§ 10. 4. 1 概述	500
§ 10. 4. 2 特征映射的基本原理	503
§ 10. 4. 3 特征映射的理论分析	514
§ 10. 5 认知机与新认知机	519
§ 10. 5. 1 增广 Hebb 规则	519
§ 10. 5. 2 认知机 (The Cognition)	520
§ 10. 5. 3 新认知机 (Neocognition)	522
§ 10. 6 混合学习方案——后向传播网络	528

第一章 总论

生命科学与工程科学的相互交叉、相互渗透、相互促进是现代科学发展的重要特点之一。神经网络(或神经计算)就是这样一门交叉学科。作为对人类智能研究的重要组成部分,它已成为神经科学、脑科学、心理学、认知科学、计算机科学、数理科学等共同关心的焦点。神经网络理论研究的问题是怎样自适应地响应于一个信息环境,自治地演化出运算能力。神经网络是信息处理的一种崭新的方法,它是对过去 40 年中一直统治着信息处理的程序化计算的第一次挑战。本书试图对神经网络的理论、实现及应用作一个较全面的介绍,侧重于神经网络的基本理论。

§ 1.1 概述

从 1946 年数字电子计算机的第一次出现到 80 年代末,基本上所有的信息处理都沿袭这样一个方法,即程序化计算,为了求解一个问题,必须设计一个算法和/或对于求解该问题的一组规则,然后正确地将它们编制成软件。

显然,只有当待完成的信息处理可以用一个已知的步骤或一组已知的规则来描述时才能够利用程序化计算;如果我们并不知道这种算法步骤和/或规则,则必须开发出来,通常这是一项艰苦耗时的工作。

对于信息处理,能否有一种新方法并不需要算法或规则而且