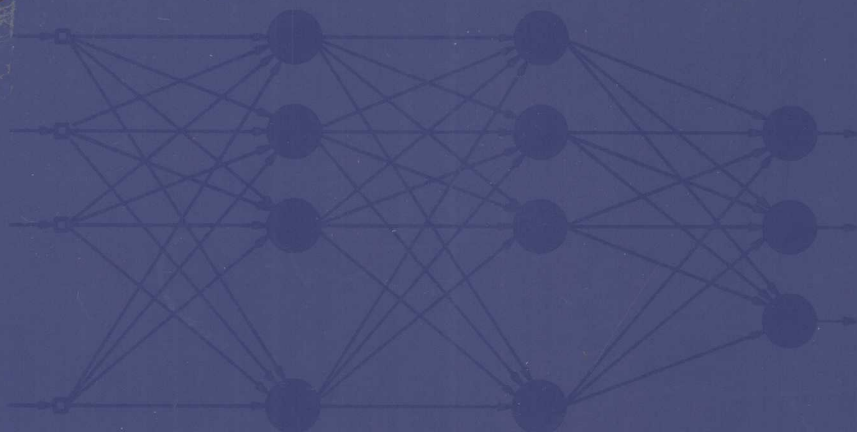


智能科学与技术

Neural Networks

神经网络

史忠植 编著



高等教育出版社
Higher Education Press

神经网络

Neural Networks

史忠植 编著

高等教育出版社

内容简介

神经网络是通过对人脑或生物神经网络的抽象和建模,研究非程序的、适应性的、大脑风格的信息处理的本质和能力。它以脑科学和认知神经科学的研究成果为基础,拓展智能信息处理的方法,为解决复杂问题和智能控制提供有效的途径,是智能科学和计算智能的重要部分。

本书系统地论述了神经网络的基本原理、方法、技术和应用,主要内容包括:神经信息处理的基本原理、感知器、反向传播网络、自组织网络、递归网络、径向基函数网络、核函数方法、神经网络集成、模糊神经网络、概率神经网络、脉冲耦合神经网络、神经场理论、神经元集群以及神经计算机。每章末附有习题,书末附有详细的参考文献

本书内容丰富,反映了当前国内外该领域的最新研究成果和动向,可作为高等院校相关专业研究生及高年级本科生的神经网络、神经计算课程的教材,也可供从事神经网络、智能信息处理、模式识别、智能控制研究与应用的科技人员参考。

图书在版编目(CIP)数据

神经网络 / 史忠植编著. —北京: 高等教育出版社, 2009.5

ISBN 978 - 7 - 04 - 026544 - 6

I. 神... II. 史... III. 人工神经元网络 IV. TP183

中国版本图书馆 CIP 数据核字(2009)第 034198 号

策划编辑 刘 英 责任编辑 刘 英 封面设计 张 志 责任绘图 郝 林
版式设计 余 杨 责任校对 杨凤玲 责任印制 陈伟光

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100120	网 址	http://www.hep.edu.cn
总 机	010-58581000		http://www.hep.com.cn
经 销	蓝色畅想图书发行有限公司	网上订购	http://www.landraco.com
印 刷	涿州市星河印刷有限公司		http://www.landraco.com.cn
		畅想教育	http://www.widedu.com
开 本	787 × 1092 1/16	版 次	2009 年 5 月第 1 版
印 张	21.25	印 次	2009 年 5 月第 1 次印刷
字 数	530 000	定 价	37.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 26544-00

前 言

神经网络是对人脑或生物神经网络的抽象和建模,研究非程序的、适应性的、大脑风格的信息处理的本质和能力。它以脑科学和认知神经科学的研究成果为基础,拓展智能信息处理的方法,为解决复杂问题和智能控制提供有效的途径,是智能科学和计算智能的重要部分。

神经网络是由大量处理单元组成的非线性大规模自适应系统,具有非线性、非局域性、非常性、非凸性等特点。神经网络在不同程度和层次上模仿人脑或生物信息处理的机理,把算法和结构一体化,具有学习能力、记忆能力、计算能力以及智能处理功能,将对智能科学和信息技术的发展产生重大影响。

自1943年麦克洛奇(McCulloch)和皮兹(Pitts)提出M-P模型之后,神经网络经过长期的研究,取得了进展,成为一种具有独特风格的信息处理方法。但目前所研究的只是基于最简单的单个神经元模型,许多艰巨而复杂的问题有待深入探讨。近年来,源自哺乳动物视觉皮层神经元信息传导模型的脉冲耦合神经网络,在解决图像处理应用问题时性能出色,引起了人们的兴趣。神经系统中信息的编码和处理在很大程度上是由大量神经元构成的集群协同活动完成的,神经元集群编码的研究正在兴起。运用现代的数学工具,从计算理论的层次研究神经网络,深入探索神经信息处理的机理。

1989年作者曾在荷兰Erasmus大学讲授神经网络课程,并编写了讲义*Introduction To Neural Networks*。1993年由电子工业出版社出版了作者编著的《神经计算》。在此基础上,吸收国内外大量最新的研究成果,包括作者在国家自然科学基金等资助下所取得的成果,写成本书,系统地阐述神经网络的基本原理和方法,反映该领域的最新进展。

全书共分4个部分,由14章组成。

(1) 神经网络基础,包括第1~8章。第1章绪论,概要介绍神经网络的发展过程、基本原理、研究内容和发展方向,并扼要介绍人脑的神经系统;第2、3章主要结合神经网络解决分类问题,重点介绍感知器、反向传播神经网络;第4章介绍具有聚类功能的自组织网络,包括自组织映射、自适应共振理论、认知器、主成分分析、独立成分分析;第5章讨论拓扑结构上闭合循环的递归网络,包括霍普菲尔特(Hopfield)模型、玻耳兹曼(Boltzmann)机;第6章讨论径向基函数神经网络,利用多维曲面对测试数据进行插值;第7章论述基于统计理论的核函数方法;第8章讨论神经网络集成,用有限个神经网络对同一个问题进行学习,以提高泛化性能。

(2) 不确定性神经网络,包括第9、10章。第9章讨论模糊神经网络,第10章介绍概率神经网络,包括贝叶斯网络、贝叶斯阴阳系统理论等。

(3) 新型神经网络,包括第11~13章。第11章讨论脉冲耦合神经网络,第12章阐述神经场理论,第13章探讨神经元集群。

(4) 神经计算机,即第14章,介绍神经元芯片和神经计算机的系统结构,给出用电子、光、分子器件等实现神经计算机的可能途径。

本书内容新颖、资料丰富,写作时力求做到由浅入深、循序渐进、条理清晰,既强调基本原理

和工程应用,又反映了国内外研究和应用的最新进展,具有科学性、先进性、实用性。在每章后面都附有一定数量的习题,并在书的最后列出了参考文献,读者可以从中进一步学习。

本书包含了作者多年的科研成果,也吸取了国内外同类教材和有关文献的精华,他们的丰硕成果和贡献是本书学术思想的重要源泉,在此谨向这些教材和文献的作者表示感谢。感谢陆汝钤、张钊、王守觉、李朝义、陈霖、郭爱克、郑南宁、何新贵、涂序彦、钟义信、徐雷、徐宗本、汪云九、罗四维、周志华等的支持和帮助,与他们的讨论给了我许多启发。本书的顺利撰写离不开中国科学院计算技术研究所智能科学实验室同事们的贡献和支持,特别是叶世伟、胡宏、张建、官秀军、何清、施智平、石志伟等;中国矿业大学丁世飞、贾伟宽、许新征、苏春阳等给予了帮助和支持。在此一并表示感谢。在本书编写和出版过程中,得到了高等教育出版社刘英编辑等的大力支持,在此致以诚挚的谢意。

本书研究工作得到自然科学基金重点项目“基于感知学习和语言认知的智能计算模型研究”(批准号:60435010)、自然科学基金项目“语义 Web 服务的逻辑基础”(批准号:60775035)的支持。感谢国家重点基础研究发展计划“基于视觉认知的非结构化信息处理理论与关键技术”(项目编号:2007CB311000)、国家 863 高技术探索项目“软件自愈与自恢复技术”(项目编号:2007AA01Z132)等项目的支持。

由于作者水平有限,加上神经网络发展很快,研究领域广泛,书中不妥和疏漏之处在所难免,恳请各位专家和广大读者不吝指教和帮助。

史忠植

2008 年 12 月于北京

目 录

第 1 章 绪论	1	敛性	44
1.1 概述	1	2.5.3 最小均方算法的评价	46
1.2 神经网络的研究历史	3	习题	46
1.3 人脑的神经系统	7	第 3 章 反向传播网络	48
1.3.1 神经元	7	3.1 概述	48
1.3.2 突触	9	3.2 反向传播网络的结构	48
1.3.3 动作电位	9	3.3 反向传播算法	50
1.3.4 离子通道	13	3.3.1 反向传播算法的基本原理	50
1.4 神经信息处理的基本原理	15	3.3.2 反向传播算法的问题	54
1.5 简单的神经网络模型	22	3.4 反向传播算法性能分析	55
1.5.1 简单线性模型	22	3.5 反向传播算法的改进	56
1.5.2 线性阈值单元	23	3.5.1 动量反向传播算法	56
1.5.3 盒中脑状态	23	3.5.2 批量更新	57
1.5.4 热力学模型	23	3.5.3 搜索然后收敛方法	57
1.6 神经网络的研究内容	25	3.5.4 自适应 BP 算法	58
1.7 神经网络的分类	28	3.5.5 共轭梯度法	58
1.8 神经网络研究的发展方向	28	3.5.6 拟牛顿法	60
习题	29	3.5.7 Levenberg - Marquardt 算法	61
第 2 章 感知器	31	3.6 反向传播网络学习程序	63
2.1 感知器的认知观点	31	习题	65
2.2 单层感知器	32	第 4 章 自组织网络	67
2.2.1 单层感知器网络结构	32	4.1 概述	67
2.2.2 感知器的学习算法	33	4.2 Kohonen 自组织映射	69
2.2.3 感知器算法的收敛性	36	4.2.1 自组织映射过程	70
2.2.4 异或问题	37	4.2.2 SOM 算法	73
2.3 多层感知器	38	4.2.3 特征映射	74
2.4 学习算法的优化	40	4.2.4 拓扑排序	76
2.4.1 最速下降法	40	4.2.5 密度匹配	77
2.4.2 牛顿方法	41	4.3 学习向量量化	79
2.4.3 高斯 - 牛顿方法	41	4.4 自适应共振理论神经网络	84
2.5 最小均方 (LMS) 算法	43	4.4.1 ART 模型的结构	85
2.5.1 最小均方算法描述	43		
2.5.2 最小均方算法的收			

4.4.2 ART 的基本工作原理	87	第 6 章 径向基函数网络	140
4.4.3 ART 模型的数学描述	92	6.1 概述	140
4.5 认知器	94	6.2 径向基函数数学基础	141
4.5.1 认知器的结构	94	6.2.1 插值计算	141
4.5.2 新认知器	98	6.2.2 模式可分性	142
4.6 主成分分析	101	6.2.3 正则化理论	143
4.6.1 基本原理	101	6.3 径向基函数网络结构	147
4.6.2 单个神经元的主成分	102	6.3.1 RBF 网络拓扑结构	147
4.6.3 单层网络主成分 提取	103	6.3.2 RBF 网络元素	148
4.6.4 侧抑制自适应主成分 提取算法	105	6.4 RBF 网络算法分析	150
4.7 独立成分分析	106	6.4.1 RBF 中心向量确定	151
4.7.1 基本概念	107	6.4.2 RBF 算法	151
4.7.2 独立成分分析神经 网络	108	6.4.3 RBF 网络性能分析	155
4.7.3 快速固定点算法	110	6.5 RBF 网络算法优化	156
习题	111	6.5.1 基于免疫算法的 RBF 网络优化	157
第 5 章 递归网络	113	6.5.2 基于遗传算法的 RBF 网络优化	158
5.1 概述	113	6.6 CMAC 网络	159
5.2 递归网络体系结构	113	6.7 泛函数连接网络	161
5.2.1 输入输出递归网络	113	6.8 小波神经网络	162
5.2.2 状态空间模型	114	6.9 过程神经网络	162
5.2.3 递归多层感知器	115	6.9.1 过程神经网络模型	163
5.2.4 二阶网络	116	6.9.2 学习算法	164
5.3 状态空间模型	117	习题	164
5.4 Hopfield 网络	118	第 7 章 核函数方法	166
5.4.1 离散 Hopfield 网络	119	7.1 概述	166
5.4.2 联想记忆	121	7.2 统计学习问题	167
5.4.3 离散 Hopfield 网络运行 程序	126	7.2.1 经验风险	167
5.4.4 连续 Hopfield 网络	130	7.2.2 VC 维	168
5.5 双向联想记忆模型	131	7.3 学习过程的一致性	168
5.6 模拟退火算法	133	7.3.1 学习一致性的经典 定义	168
5.7 玻尔兹曼机	135	7.3.2 学习理论的重要定理	169
5.7.1 网络结构	135	7.3.3 VC 熵	169
5.7.2 学习算法	136	7.4 结构风险最小归纳原理	170
习题	138	7.5 支持向量机	172
		7.5.1 线性可分	173

7.5.2 线性不可分	174	10.5 密度估计的一致性	205
7.6 核函数	176	10.6 概率神经网络	206
7.6.1 多项式核函数	176	10.7 激活函数	208
7.6.2 径向基函数	176	10.8 贝叶斯阴阳系统理论	209
7.6.3 多层感知器	176	习题	211
7.6.4 动态核函数	176	第 11 章 脉冲耦合神经网络	212
7.7 核主成分分析	178	11.1 概述	212
习题	179	11.2 视觉皮层理论	212
第 8 章 神经网络集成	180	11.2.1 Hodgkin - Huxley 模型	213
8.1 概述	180	11.2.2 FitzHugh - Nagumo 模型	214
8.2 神经网络集成的基本原理	181	11.2.3 Eckhorn 模型	214
8.3 神经网络集成的方法	181	11.3 脉冲耦合神经网络模型	215
8.4 结论生成方法	182	11.4 交叉皮层模型	217
8.5 个体生成方法	183	11.5 贝叶斯连接域神经网络 模型	217
8.5.1 Boosting 算法	184	11.5.1 带噪声的神经元发放 方式	217
8.5.2 Bagging 算法	185	11.5.2 神经元输入的贝叶斯 耦合方式	218
8.6 基于 Bagging 的聚类	186	11.5.3 神经元之间的竞争 关系	219
8.7 神经网络集成系统的规则 获取	187	11.6 贝叶斯连接域神经网络模型 在特征捆绑中的应用	220
8.8 神经专家系统	189	习题	224
8.8.1 知识表示的神经网络 方法	189	第 12 章 神经场理论	225
8.8.2 推理机制	190	12.1 概述	225
习题	192	12.2 信息几何	226
第 9 章 模糊神经网络	193	12.2.1 微分流形	226
9.1 概述	193	12.2.2 切向量和切向量 空间	228
9.2 算术模糊神经网络	194	12.2.3 Riemanniann 流形	229
9.3 模糊逻辑	195	12.2.4 仿射联络	229
9.4 模糊联想记忆	196	12.2.5 测地线	230
9.5 神经模糊推理系统	198	12.2.6 Levi - Civita 曲率	231
9.6 神经网络近似逻辑	201	12.2.7 流形上向量的平移 变换与平坦流形	231
习题	202		
第 10 章 概率神经网络	203		
10.1 概述	203		
10.2 贝叶斯定理	203		
10.3 概率密度函数	204		
10.4 模式分类的贝叶斯判定 策略	205		

12.3 统计流形上的 Riemann 度量	232	13.5.6 模拟结果	261
12.3.1 参数分布族的几何	232	13.6 集群编码	264
12.3.2 切空间及其统计表示	233	13.6.1 运动方向的神经元 集群编码	265
12.3.3 Riemann 度量和 Fisher 信息	233	13.6.2 神经编码的精确性	265
12.4 神经场理论模型	234	13.6.3 Bayesian 方法的最佳译码	266
12.4.1 神经场表示	234	13.7 神经元集群时空编码	267
12.4.2 神经场学习理论	236	习题	268
12.5 基于 Fisher 分的朴素贝叶斯分类器	241	第 14 章 神经计算机	269
12.5.1 Fisher 分	241	14.1 神经计算机的体系结构	269
12.5.2 基于 Fisher 分的朴素贝叶斯分类器构建算法	242	14.1.1 全硬件实现	269
12.6 动态神经场模型	243	14.1.2 虚拟实现	271
12.6.1 数学框架	243	14.2 电子神经器件	272
12.6.2 动态行为	244	14.2.1 数字神经芯片	273
12.6.3 图案形成	244	14.2.2 电压模式神经器件	275
12.6.4 行波	245	14.2.3 开关电容神经网络	276
习题	246	14.2.4 电流模式神经器件	276
第 13 章 神经元集群	247	14.3 电子神经计算机	277
13.1 概述	247	14.3.1 神经网络协处理器	278
13.2 大脑皮层	248	14.3.2 并行处理机阵列	282
13.3 嗅觉神经系统的 K 系列模型	251	14.3.3 脉动神经计算机	283
13.3.1 KO 模型	252	14.4 基于细胞自动机的人工脑	284
13.3.2 KI 模型	252	14.5 光神经计算机	287
13.3.3 KII 模型	253	14.5.1 矩阵处理器	287
13.3.4 KIII 模型	253	14.5.2 空间光调制器	289
13.4 皮层功能柱的细胞结构	255	14.5.3 光互连	290
13.5 皮层功能柱理论模型	257	14.5.4 光全息存储器	290
13.5.1 神经元模型	258	14.5.5 全光神经计算机	290
13.5.2 突触模型	260	14.6 光电神经计算机	291
13.5.3 网络结构	260	14.7 分子神经计算机	293
13.5.4 网络的输入	260	14.7.1 分子计算的宏-微模式	293
13.5.5 度量指标和计算方法	261	14.7.2 生物芯片	294
		14.7.3 分子神经计算机的体系结构	294
		习题	297
		参考文献	298

第 1 章 绪 论

神经网络通过对人脑或生物神经网络基本特性的抽象和建模,研究非程序的、适应性的、大脑风格的信息处理的本质和能力。本章主要介绍神经网络的研究历史、信息处理的基本原理、研究内容、研究的发展方向等问题。

1.1 概述

神经网络(Neural Networks, NN),也称做人工神经网络(Artificial Neural Networks, ANN),或神经计算(Neural Computing, NC),是对人脑或生物神经网络的抽象和建模,具有从环境学习的能力,以类似生物的交互方式适应环境。神经网络是智能科学和计算智能的重要部分,以脑科学和认知神经科学的研究成果为基础,拓展智能信息处理的方法,为解决复杂问题和实现自动控制提供有效的途径。

从 1946 年第一台电子数字计算机(ENIAC)问世以来直至今,大多数信息处理都采用程序式计算方式。这种方式解题需要设计算法或规则,并正确地编制成软件,然后才能进行问题求解。这种解题方式必须考虑 3 个因素:

- ① 问题的形式化;
- ② 可进行计算的算法;
- ③ 计算的复杂性。

比较计算机和人的处理能力,其差别是惊人的。一方面,一个人能很容易识别面孔,理解语言,而一台巨型机却很难识别出一棵树来。另一方面,用计算机进行计算,可以很快得到答案,其计算能力大大超过了人。那么数字计算和辨识物体之间究竟有哪些差别呢?

辨识物体是不能简单地加以定义的。要识别一棵树,就必须给出树的全部定义。做出这样一种定义,等于要描述树的每一个可以想象到的变量。这类问题构成了随机问题。所谓随机问题,就是那些需要具备某一系统的实际上每种可能状态的知识才能解答的问题。因此,为解决一个随机问题,就要求记忆所有可能的解答,当给定输入数据时,从所有可能的解答的集合中迅速选出最合适的答案。而像数学一类的计算问题,其解答通常可以用一种算法简洁地表示出来,也就是说,可以用一个精确的指令系列来表示,该指令系列规定了如何处理输入数据以得到答案。

信息处理的一种新方法并不需要开发算法和规则,极大地减少了软件的工作量,这种方法称为神经网络。神经网络是一门崭新的信息处理学科,它从神经生理学和认知科学研究成果出发,

应用数学方法研究并行分布的、非程序的、适应性的、大脑风格的信息处理的本质和能力。神经网络中主要的信息处理结构是人工神经网络。

神经信息处理是介于常规处理形式和人脑处理形式的中间处理形式。一方面,神经网络企图模仿人脑的功能,而另一方面许多实现技术又是常规的。表 1-1 给出了这 3 种信息处理范型的主要特点。神经信息处理许多特性与人脑相似,诸如联想、概括、并行搜索、学习和灵活性。

表 1-1 3 种信息处理范型

项目	常规计算机	神经计算	人脑
信息表示	指令 + 数据	网络连接处理单元功能 + 权值	内部神经连接
程序设计	指令 + 初始数据	网络拓扑 + 模式训练	模式训练
处理	数字	数字或模拟	模拟
体系结构	$1 \sim 10^4$ 处理器	$1 \sim 10^6$ 处理器	10^{11} 神经元
硬件	集成电路	电、光、生物器件	神经元
开关速度	1 ms	1 ns ~ 1 ms	1 ms
工艺	硅器件	硅器件、光、分子	生物的

人脑信息处理的特点如下:

① 大规模并行处理。人脑神经元之间传递神经冲动是以毫秒计的,比普通的电子计算机慢得多。但人们通常能在 1 ms 内对外界事物做出判断和决策。这对传统的计算机或人工智能是做不到的。由此可知,人脑的“计算”必定是建立在大规模并行处理的基础上。人善于在复杂环境中做出判断,从整体上识别事物。神经网络的大规模并行处理与多处理机构成的并行系统是不同的。

② 具有很强的容错性,善于联想、概括、类比和推广。每天有大量神经细胞正常死亡,但不影响大脑正常的功能;大脑局部损伤会引起某些功能衰退,但不是功能突然丧失。在计算机中,元器件的局部损坏,或者程序中的微小错误都可能引起严重的后果,即表现出极大的脆弱性。人脑与计算机信息处理的巨大差别在于对信息的记忆和处理方式不同。计算机的模式是信息局部存储,按程序提取有关的信息,送到运算器处理。大脑中信息的记忆,特别是长期记忆是通过改变突触的效能实现的,即信息存储在神经元间连接强度的分布上,信息的记忆和处理是合二为一的。这一点,神经网络与大脑信息处理方式极其相似。

③ 具有很强的自适应能力。人脑功能受先天因素制约,但后天因素,如经历、训练、学习等也起重要作用。这表明人脑具有很强的自适应性和自组织性。神经网络与符号处理不同,前者强调系统的自适应或学习过程,同一网络因学习方法及内容不同,可具有不同的功能;符号处理强调程序编写,系统的功能取决于编写者的知识和能力。

由上可知,脑是最复杂、最完美、最有效的一种信息处理装置,人们正以极大的兴趣研究它的结构和机理。这种研究与 20 世纪初的物理学和 20 世纪 50 年代的分子生物学一样,正酝酿着重大的突破,而这一突破将给整个科学的发展带来巨大而深远的影响。人们对大脑的认识已深入到探索脑的核心问题,鉴定出了一系列涉及脑工作的重要分子,在感知、行为、学习和记忆方面都

取得了重要进展。这表明人们将有可能最终揭开大脑这个人体最复杂系统的奥秘,为现代科技发展寻找新的道路。借助大脑工作原理,有可能使信息处理获得新的突破。

正因为如此,神经科学受到世界各发达国家的高度重视。美国国会通过决议将1990年1月5日开始的10年定为“脑的十年”。国际脑研究组织号召它的成员国将“脑的十年”变为全球行动。美国国防部高级研究计划局(DARPA)制定的8年研究计划中,神经网络是重要的方向。1986年日本政府提出了“人类前沿科学计划”(HFSP)研究计划,1992年提出“真实世界计算”(RWC)研究计划。德国从1988年开始执行“神经信息论”的研究计划。

脑科学、神经生理学、病理学主要研究神经网络的生理机理,如神经元、突触、化学递质、脑组织等的构成和工作过程。而认知科学、计算机科学主要探索人脑信息处理的微结构理论,寻求新的途径,解决当前计算机和传统人工智能难以处理的问题。以此为背景,以人工神经网络为基础,形成了神经网络的新学科。

目前,对大脑思维的过程了解仍然很肤浅,人工神经网络模拟的研究还很不充分,人们面临的是一个充满未知的新领域。神经网络将在基本原理方面进行更深刻的探索。

神经网络的发展与神经科学、认知科学、计算机科学、人工智能、信息科学、机器人学、微电子学、光计算、分子生物学等有关,是一门新兴的边缘交叉学科。神经网络研究的主要目标如下:

① 理解脑系统为何具有智能。这些计算与符号表示的形式操作处理不同,人脑是如何组织和实施这些“计算”的。

② 研究各种强调“计算能力”的神经网络模型,并不着重于这些模型的生物学保真程度。

③ 研究大规模并行自适应处理的机理。

④ 研究神经计算机的体系结构和实现技术。

1.2 神经网络的研究历史

20世纪40年代中期,在科技发展史上出现了两种新的计算工具和一种描述神经网络工作的数学模型。以运算放大器为基本部件的模拟计算机,对于求解微分方程特别简单易行,因此在自动控制等领域得到应用。但是,由于实现的困难,对电学模拟量进行运算的模拟机在精度上难以提高。所以模拟计算机只有一个短暂的发展期,随后就很少采用了。而以双稳态触发器为主要运算器件、以布尔代数作为逻辑基础、以中央处理器为控制核心的数字计算机,在与模拟计算机的竞争中逐渐取得上风。原因在于数字计算机的通用性强、精度高,而且随着程序设计语言和软件的发展,使用和解决具体问题的能力日益增强。由于大规模集成电路的发展,数字计算机的体积缩小、速度提高、成本价格下降。数字计算机系统的面貌日新月异,新产品不断推向市场。

与数字计算机发展的一帆风顺相比,神经网络的发展经历了一条曲折的道路。

1. 麦克洛奇—皮兹(M-P)神经网络模型

现在神经网络的理论准备可追溯到20世纪40年代。当时一些神经科学家、生理学家对人脑神经元的电生理方面做了不少工作。1943年,心理学家麦克洛奇(McCulloch)和数理逻辑学家皮兹(Pitts)在《数学生物物理公报》(*Bulletin of Mathematical Biophysics*)上发表了关于神经网络的数学模型。这个模型,现在一般称为M-P神经网络模型^[314]。他们总结了神经元的一些基本生理特性,提出神经元形式化的数学描述和网络的结构方法,从此开创了神经网络的时代。

尽管当时年轻的数学家皮兹对于能够用数理逻辑工具描述神经网络的活动规律,欣喜若狂,以致耽误了他取得博士学位,但整个科学界并未予以重视,而认为这仅是一种纯理论上的探索。M-P模型中的形式神经元的性能是简单的,可是网络的计算能力是巨大的。这种巨大的计算潜力来源于神经元之间丰富的联系和整个网络的平行计算。因此,可以认为M-P模型是现代连接主义学派的一个原型。但是,当时的计算机界并不真想构造一个电子模型来实现M-P网络。注意实验的神经科学家也不理解M-P模型的理论价值。只有数学家克里纳(Kleene)把M-P模型整理清楚,抽象发展成一种有限自动机理论。

2. 赫布规则

神经计算机与传统计算机的最大区别之一是它能像动物那样进行学习。动物之所以能学习,原因在于它有日益进化的神经系统。但是,这个学习过程是如何在神经系统中进行的,至今还没有一个明确的结论。1949年心理学家赫布(Hebb)提出突触连接强度可变的假设^[205]。这个假设认为学习过程最终发生在神经元之间的突触部位,突触的连接强度随着突触前后神经元的活动而变化。现在,这个规则一般称为赫布规则。赫布提出的学习规则为神经网络的学习算法奠定了基础。现在多数的学习规则遵循这一规则。赫布在《行为的组织》一书中还阐明了另外两个关于神经网络功能的重要论点:一是神经细胞集合的概念,神经元相互连接成可自强化的小集团,从而形成神经系统中信息的表象;二是表象的分布性,即网络中许多细胞参与代表某种事物的表象。

3. 感知器模型

1957年罗森勃拉特(F. Rosenblatt)提出的感知器(Perceptron)模型是该领域的一个重要进展,第一次把神经网络研究从纯理论的探讨付诸工程实现,掀起神经网络研究的第一次高潮。感知器由阈值神经元构成,分成若干层神经网络。第1层为视网膜层,接受来自外界的输入,然后投射到第2层。第2层以随机方式连到第3层。最后一层是反应层。罗森勃拉特提出的感知器模型中已经包含了许多现代神经计算机的基本原理。它是一种学习和自组织的心理学模型,整个模型的结构又大体上符合神经生理学知识。模型的学习环境有噪声,内部结构又有随机联系。这些情况符合动物学习的自然环境。感知器中的学习规则是突触强化律。这种机器显然有可能用于模式识别、联想记忆等方面。因此,当时世界上有许多实验室纷纷仿效,设计多种多样的电子装置进行声音识别、文字识别和学习记忆研究。据统计,当时有上百个实验室在研究这类机器。1962年威德罗(Widrow)提出了自适应线性元件(Adaline),它是连续取值的线性网络,主要用于自适应系统。这与当时处于主导地位的以符号推理为特征的传统人工智能途径完全不同,因而形成了神经网络、脑模型研究的高潮。

20世纪60年代以后,由于数字计算机的飞速发展,传统人工智能初期研究的巨大成功,吸引了大量专家投入数字计算机的研究。同时发现神经网络极大的局限性。明斯基(Minsky)和佩珀特(Papert)仔细分析了以感知器为代表的神经网络系统的功能及局限后,于1969年出版了*Perceptrons*一书,指出感知器仅能解决一阶谓词逻辑,不能解决高阶谓词问题,对神经网络的功能持悲观态度^[324]。他们的论点极大地影响了对神经网络的研究,使较多的人转去研究当时发展较快的以逻辑为基础的人工智能和知识工程。加上冯·诺伊曼(von Neumann)串行计算机在技术上、规模上、速度上都发展很快,从而掩盖了发展新型计算机和人工智能新途径的必要性和迫切性,致使人工神经网络的研究进入低潮。

4. 霍普菲尔特神经网络模型

一方面,20世纪70年代末以来,随着人工智能在模拟人的某些认知活动取得很大的进展,专家系统、智能计算机受到重视的同时,人们特别感到了传统的人工智能系统与人的自然智能相比存在一些明显的不足。人工智能在感知能力上同人的差距很大。人能够毫不费力地识别各种复杂的事物,能从记忆的大量信息中迅速找到需要的信息,人具有自适应、自学习等创新知识的能力。这些都是现有计算机无法比拟的。因此人们又重新将目标转向神经网络的研究上,试图通过对人脑神经系统的结构、信息加工、记忆和学习机制的分析、探索,提出解决上述差距的新思想、新方法。另一方面,学术界对于复杂系统的研究取得了许多进展。普里高京(Prigogine)提出非平衡系统的自组织理论,即耗散结构理论,获得诺贝尔奖。哈肯(Haken)研究大量元件联合行动而产生的有序的宏观表现,创立了协同学(Synergetics)。近年来广泛研究的混沌(chaos)动力学和奇异吸引子理论揭示了复杂系统行为。这些工作,从抽象意义上讲,都是研究复杂行为系统如何通过元件之间的相互作用,系统的结构上由无序到有序,功能上由简单到复杂,类似于生物系统的进化和自组织过程,以及认知系统的学习过程。与此同时,神经科学和脑科学日益受到人们的重视,在感觉系统,特别是视觉研究发现的侧抑制原理、感觉野概念、皮层的功能柱结构,以及信息处理的平行、层次观点,被证明是神经系统处理信息的普遍原则。芬兰电子工程师科霍南(Kohonen)提出了自组织映射模型、日本NHK的福岛邦彦(Fukushima)提出的认知机、美国波士顿大学的格罗斯伯格(Grossberg)关于感知觉的共振适应理论、日本甘利俊一(S. Amari)关于神经网络有关数学理论的研究、安德森(Anderson)提出了盒中脑(Brain State in Box,BSB)模型,这些都是20世纪70年代到20世纪80年代初进行的工作。

20世纪80年代以来,有关神经网络的研究进展非常迅速。1982年美国加州工学院物理学家霍普菲尔特(Hopfield)的工作被称为是突破性的^[219]。他提出了离散的神经网络模型,从而有力地推动了神经网络的研究,标志着神经网络研究高潮的又一次到来。他引入李雅普诺夫(Lyapunov)函数(叫做“计算能量函数”),给出了网络稳定判据。1984年霍普菲尔特又提出了连续神经网络模型,其中神经元动态方程可以用运算放大器来实现,因此神经网络可以用电子线路来仿真^[220]。它与VLSI有直接对应关系,为神经计算机的研制奠定了基础。同时它还可用于联想记忆和优化计算,开拓了神经网络用于计算机的新途径。次年,美国加州理工学院和贝尔实验室合作制成256个神经元和64个可编程的人工神经元。前者由25000个晶体管和10万个电阻集成在1/4平方英寸^①的芯片上。

5. 并行分布处理理论

费尔德曼(Feldmann)和巴拉德(Ballard)的连接网络模型指出了传统的人工智能计算与生物“计算”的区别,提出了并行分布处理的计算原则。1985年,欣顿(Hinton)和塞杰诺斯基(Sejnowsky)提出了一个可行的算法,称为玻耳兹曼(Boltzmann)机模型。他们借用了统计物理学的概念和方法,首次提出了多层网络的学习算法。1986年,鲁梅尔哈特(Rumelhart)和麦克莱伦德(McClelland)等人提出并行分布处理(Parallel Distributed Processing,PDP)的理论^[397]。致力于认知的微观结构的探索,并提出了多层网络的误差传播学习法,或称反传(Back-Propagation,BP)算法。这种算法根据学习的误差大小,从后向前修正各层次之间的连接权值。通过不断的学习

^① 1英寸=2.54厘米。

和修正,可以使网络的学习误差达到最小。反传算法从实践上证明神经网络的运算能力很强,可以解决许多具体问题。神经计算机的先驱赫希特-尼尔逊(Hecht-Nielsen)提出了对传(Counter-Propagation)网络模型,可用于图像压缩和统计分析等。1988年美国加州大学的蔡少堂(L. O. Chua)等人提出了细胞神经网络模型,与一般神经网络一样,它是一个大规模非线性仿真系统,同时具有细胞自动机的动力学特征。

6. 神经计算机的研制

神经计算机实现的途径有半导体电子器件、光学元件和分子器件,近年来的研究取得了一定进展。目前,大量的工作集中于硅半导体 VLSI 电路制作,利用 CMOS 工艺、数字和模拟混合系统来实现。美国 R. Hecht-Nielsen 创建的 Hecht-Nielsen 神经计算机公司已经开发了两代产品。美国 IBM 公司推出的神经网络工作站也已进入市场。1990年1月上旬,日本富士通公司研制成功每秒运算5亿次的神经计算机,是当时世界上运算速度最快的神经计算机。它由256个神经细胞采用化合物半导体高集成化技术研制成世界上第一个能识别26个字母的光神经电路芯片,具有记忆、联想和判断功能。日本自动翻译电话研究所与美国卡内基-梅隆大学合作研制了一种具有英、日文自动翻译功能的新型神经计算机。该机由模仿人脑构造、连接相当于右脑和左脑的两个神经块组成,能从事简单的翻译工作。表1-2给出了神经网络的研究历史。

表 1-2 神经网络的研究历史

时间/年	贡 献 者	成 果 名 称
1943	McCulloch, Pitts	M-P 模型
1949	Hebb	赫布规则
1951	Minsky	Snark 神经计算机
1957	Rosenblatt	感知器模型
1962	Widrow	自适应线性元件
1969	Minsky, Papert	著作 <i>Perceptrons</i>
1972	Fukushima	认知器模型
1974	Werbos	反向传播模型
1977	Anderson	盒中脑(BSB)模型
1978—1990	Grossberg	自适应共振理论(ART)
1980	Kohonen	自组织映射
1982	Hopfield	Hopfield 离散神经网络模型
1984	Hopfield	Hopfield 连续神经网络模型
1985	Hinton, Sejnowsky	玻耳兹曼机
1986	Rumelhart, McClelland	PDP 理论
1986	Hecht-Nielsen	对传网络模型
1986	Marks II	交替投影模型
1988	L. O. Chua	细胞神经网络模型

自 20 世纪 80 年代中期以来,世界上许多国家都掀起了神经网络的研究热潮。从 1985 年开始,专门讨论神经网络的学术会议规模逐步扩大。1987 年在美国召开了第一届神经网络国际会议,并发起成立国际神经网络学会(INNS)。为了推动神经网络的研究,出版了几种专门的学术刊物,如 *Neural Networks*、*Connection Science*、*IEEE Transactions on Neural Networks*、*Neural Computation* 等。

在国际研究潮流的推动下,我国在神经网络这个新兴的研究领域取得了一些研究成果,几年来形成了一支多学科的研究队伍,组织了不同层次的讨论会。1986 年中国科学院召开了“脑工作原理讨论会”,1989 年 5 月在北京大学召开了“识别和学习国际学术讨论会”,1990 年 10 月中国自动化学会、中国计算机学会、中国心理学会、中国电子学会、中国生物物理学会、中国自动化学会、中国物理学会、中国通信学会等 8 个学会联合召开“中国神经网络首届学术大会”,论文内容涉及脑功能及生物神经网络模型、神经生理与认知心理模型、人工神经网络模型、神经网络理论、新的学习算法、神经计算机、VLSI 及光学实现、联想记忆、神经网络与人工智能、神经网络与信息处理、神经网络与模式识别、神经网络与自动控制、神经网络与组合优化、神经网络与通信等。1992 年 11 月,国际神经网络学会、IEEE 神经网络学会、中国神经网络学会等联合在北京召开了神经网络国际会议。为了培养神经网络方面的研究人才,不少高等院校开设了“神经网络”、“人工神经网络”、“神经计算”及其有关的课程。许多单位开展了神经网络和神经科学的基础研究和应用开发,取得了一定进展。

1.3 人脑的神经系统

人脑神经系统的主要细胞组成是神经细胞和神经胶质细胞。神经系统表现出来的一切兴奋、传导和整合等机能特性都是神经细胞的机能。胶质细胞占脑容积一半以上,数量大大超过了神经细胞,但在机能上只起辅助作用。

1.3.1 神经元

神经细胞是构成神经系统最基本的单位,故通称为神经元。一般包括神经细胞体(soma)、树突(dendrites)和轴突(axon)三部分(神经元的突起是神经元胞体的延伸部分,由于形态结构和功能的不同,可分为树突和轴突)。神经元的一般结构如图 1-1 所示。

1. 细胞体

细胞体(简称胞体),是神经元的主体,位于脑和脊髓的灰质及神经节内,其形态各异,常见的形态为星形、锥形体、梨形和圆球形等。胞体大小不一,直径在 5 ~ 150 μm 之间。胞体是神经元的代谢和营养中心。胞体的结构与一般细胞相似,有核仁、细胞膜、细胞质和细胞核。胞内原浆在活细胞内呈颗粒状,经固定染色后显示内含神经原纤维、核外染色质(尼氏体、高尔基氏体、内质网和线粒体等)。神经原纤维是神经元特有的。

(1) 细胞膜

胞体的细胞膜(简称胞膜)和突起表面的膜,是连续完整的细胞膜。除突触部位的胞膜有特优的结构外,大部分胞膜为单位膜结构。神经细胞膜的特点是一个敏感而易兴奋的膜。在膜上有各种受体(receptor)和离子通道(ionic channel),二者各由不同的膜蛋白所构成。形成突触部分

的细胞膜增厚。膜上受体可与相应的化学物质神经递质结合。当受体与乙酰胆碱递质或 γ -氨基丁酸递质结合时,膜的离子通透性及膜内外电位差发生改变,细胞膜产生相应的生理活动——兴奋或抑制。

(2) 细胞核

细胞核(简称胞核)多位于神经细胞体中央,大而圆,异染色质少,多位于核膜内侧,常染色质多,散于核的中部,故着色浅,核仁 1~2 个,大而明显。细胞变性时,核多移向周边而偏位。

(3) 细胞质

细胞质(简称胞质)位于核的周围,又称核周体(perikaryon),其中含有发达的高尔基复合体、滑面内质网,丰富的线粒体、尼氏体及神经原纤维,还含有溶酶体、脂褐素等结构。具有分泌功能的神经元,胞质内还含有分泌颗粒,如位于下丘脑的一些神经元。

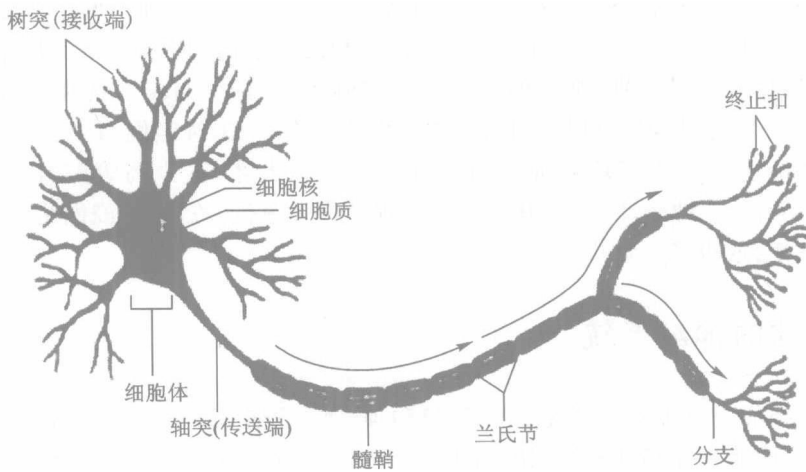


图 1-1 神经元的结构

2. 树突

树突是从胞体发出的一个或多个突起,呈放射状。胞体起始部分较粗,经反复分支而变细,形如树枝状。树突的结构与胞体相似,胞质内含有尼氏体、线粒体和平行排列的神经原纤维等,但无高尔基复合体。在特殊银染标本上,树突表面可见许多棘状突起,长约 0.5~1.0 μm ,粗约 0.5~2.0 μm ,称树突棘(dendritic spine),是形成突触的部位。一般电镜下,树突棘内含有数个扁平的囊泡称棘器(spine apparatus)。树突的分支和树突棘可扩大神经元接受刺激的表面积。树突具有接受刺激并将冲动传入细胞体的功能。

3. 轴突

轴突每个神经元只有一根,在胞体上发出的轴突多呈锥形,称轴丘(axon hillock),其中没有尼氏体,主要有神经原纤维分布。轴突自胞体伸出后,开始的一段称为起始段,长约 15~25 μm ,通常较树突细,粗细均匀,表面光滑,分支较少,无髓鞘包卷。离开胞体一定距离后,有髓鞘包卷,即为有髓神经纤维。轴突末端多呈纤细分支称轴突终末(axon terminal),与其他神经元或效应细胞接触。轴突表面的细胞膜称轴膜(axolemma),轴突内的胞质称轴质(axoplasm)或轴浆。轴质内有许多与轴突长轴平行的神经原纤维和细长的线粒体,但无尼氏体和高尔基复合体,因此,轴突内不能合成蛋白质。轴突成分代谢更新以及突触小泡内神经递质,均在胞体内合成,通过轴