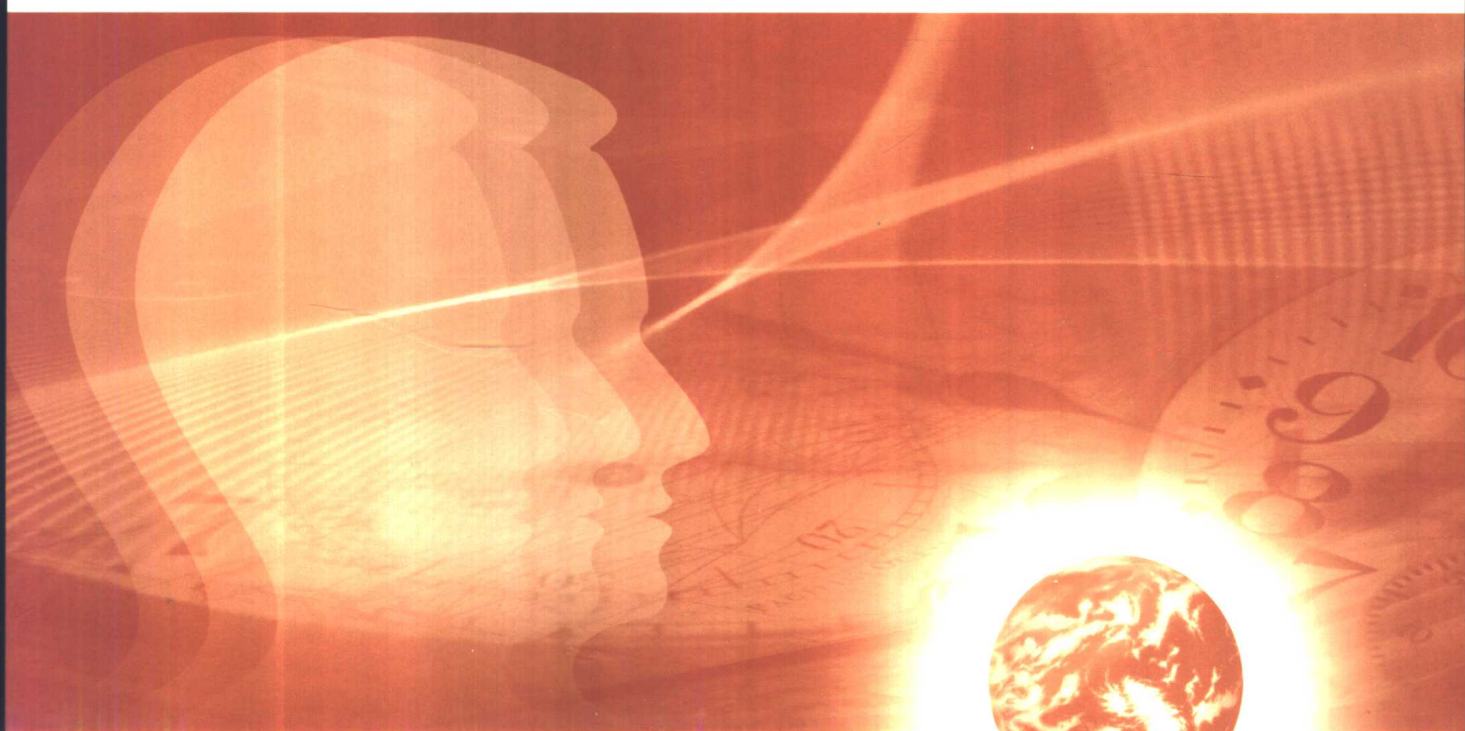


高等院校计算机科学与技术
“十五”规划教材

人工神经网络原理 及仿真实例



●
高隽
编著

 机械工业出版社
CHINA MACHINE PRESS



高等院校计算机科学与技术“十五”规划教材

人工神经网络原理及仿真实例

高 隽 编著



机械工业出版社

本书以神经网络结构为主线,以学习算法为副线,详细介绍了神经网络结构和算法步骤,并给出实例和练习,目的是使读者易看懂,能动手,会应用。主要内容包括:人工神经网络简介、单层前向网络及 LMS 学习算法、多层前向网络及 BP 学习算法、支持向量机及其学习算法、Hopfield 神经网络与联想记忆、随机神经网络及模拟退火算法、竞争神经网络和协同神经网络。每章均给出了基于 MATLAB 的仿真实例以及练习。

本书可作为电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等专业研究生和高年级本科生的教材,同时对有关专业领域的研究人员和工程技术人员也有重要的参考价值。

图书在版编目(CIP)数据

人工神经网络原理及仿真实例/高隽编著. —北京:机械工业出版社, 2003.7
高等院校计算机科学与技术“十五”规划教材

ISBN 7-111-12591-6

I. 人... II. 高... III. ①人工神经网络—高等学校—教材②计算机仿真—高等学校—教材 IV. TP183②TP391.9

中国版本图书馆 CIP 数据核字(2003)第 056288 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

策 划: 胡毓坚

责任编辑: 时 静

责任印制: 付方敏

三河市宏达印刷有限公司印刷·新华书店北京发行所发行

2003 年 8 月第 1 版·第 1 次印刷

787mm×1092mm 1/16·13.75 印张·335 千字

0001—5000 册

定价: 20.00 元

凡购本图书,如有缺页、倒页、脱页,由本社发行部调换

本社购书热线电话(010) 68993821、88379646

封面无防伪标均为盗版

出版说明

信息技术高度普及的今天,具备一定层次的信息技术素养成为社会素质教育的一个重要目标,由此对高等院校的计算机专业教育提出了更高更新的要求。教育水平提高的关键是教学质量,那么对教学质量有直接影响的教材建设就成为了计算机专业教育的根本,为重中之重。

适逢高等院校计算机专业教育改革的关键时期,为配合相关的教材建设,机械工业出版社会同全国在该领域内享誉盛名、具备雄厚师资和技术力量的高等院校,包括清华大学、上海交通大学、南京大学、成都电子科技大学、东南大学、西安电子科技大学、解放军理工大学、北京科技大学等重点名校,组织了多位长期从事教学工作的骨干教师,集思广益,对当前高等院校的教学现状开展了广泛而深入的研讨,继而紧密结合当前技术发展需要并针对教学改革所提出的问题,精心编写了这套面向普通高等院校计算机专业的系列教材,并陆续出版。

本套教材内容覆盖了普通高等院校计算机专业学生的必修课程,另外还恰如其分地添加了一些选修课程,总体上分为基础、软件、硬件、网络和多媒体五大类。在编写过程中,对教学改革力度比较大、内容新颖以及各院校急需的并且适应社会经济发展的新教材,优先选择出版。

本套教材注重系统性、普及性和实用性,力求达到专业基础课教材概念清晰、深度合理的标准,并且注意与专业课教学的衔接;专业课教材覆盖面广、深浅适中,在体现相关领域最新发展的同时注重理论联系实际。全套教材体现了教育改革的最新思想,可作为高等院校计算机科学与技术专业的教学用书,同时也是培训班和自学使用的最佳教材。

机械工业出版社

前 言

人工神经网络是近年来的热点研究领域,涉及到电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等诸多学科,其应用领域包括:建模、时间序列分析、模式识别和控制等,并在不断的拓展。

本书的总体编写思路为:介绍背景和历史,说清理论和方法,详细介绍网络结构和算法步骤,并给出实例和练习,使读者易看懂,能动手,会应用。基于这一编写思路,本书以网络结构为主线,以学习算法为副线,介绍前向神经网络、反馈神经网络、随机神经网络和竞争神经网络,并介绍相应的学习算法,给出基于 MATLAB 的仿真实例。

第 1 章引言对人工神经网络进行简介。第 2 章首先介绍单层感知器及其学习算法,然后介绍自适应线性元件以及 LMS 算法,最后结合讨论的问题给出典型算法的仿真实例。第 3 章介绍多层前向神经网络中的多层感知器和径向基网络,重点介绍了 BP 学习算法的原理、步骤以及算法的改进,并将多层感知器和径向基网络进行了比较;统计学习理论和支持向量机具有很好的性能,是近年来的研究热点。虽然从神经网络结构划分上可将其看成多层前向神经网络,但考虑到其重要性,本书在第 4 章专门介绍了统计学习理论和支持向量机,首先介绍统计学习理论的基本内容,在此基础上,阐述线性支持向量机和非线性支持向量机的机理,并将其与多层前向网络相比较;接着,介绍用支持向量机如何解决分类问题及其相应的学习算法;最后,以几个实验为例,更为详细地阐述了支持向量机的工作过程及其结果。第 5 章介绍反馈型神经网络,介绍了神经动力学基本理论,重点介绍著名的离散 Hopfield 网络和连续 Hopfield 网络模型,再通过联想记忆问题,介绍 Hopfield 网络权值的设计方法,而 Hopfield 网络用于最优化计算,通过仿真实例加以介绍。第 6 章介绍随机型神经网络,重点介绍了 Boltzmann 机的原理、运行步骤、学习规则及其算法的改进,并重点介绍了模拟退火算法及其在组合优化问题上的应用。第 7 章介绍竞争型神经网络,主要介绍 Hamming 网、自组织映射网络、学习矢量量化以及主分量分析。通过对网络结构的不断改进,充分展示竞争学习的主要思想。作为具有深厚数理基础的另一种竞争型神经网络——协同神经网络,在第 8 章专门介绍。第 8 章首先介绍协同学的基本概念和数学模型,然后在协同学基本方程的基础上,介绍协同神经网络的数学模型,并详细介绍其网络结构、运行步骤及其学习算法,最后给出仿真实例。

本书可作为电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等专业研究生和高年级本科生的教材,同时对有关专业领域的研究人员和工程技术人员也有重要的参考价值。

由于本人水平有限,另外人工神经网络本身也在不断地丰富和发展,书中不妥和错误之处在所难免,恳请广大读者批评指正。

高 隽

致 谢

本书是在国家自然科学基金青年基金项目、国家自然科学基金面上项目、安徽省自然科学基金项目、教育部高校骨干教师资助计划项目和教育部优秀青年教师资助计划项目研究工作的基础上完成的。在此，特向国家自然科学基金委员会、安徽省自然科学基金委员会和教育部表示衷心的感谢！

本书编写过程中参考了大量的学术专著和论文，由于所参考的学术论文过多，无法一一列举，对此特向未列出的参文作者表示歉意！向从事人工神经网络研究的前辈专家、老师和同仁表示由衷的敬意和感谢！

合肥工业大学“图像信息处理研究室”是我事业的基石，所取得的点点滴滴成绩都离不开图像室同仁们的支持和帮助。在此，向鲍捷、胡良梅、梁栋、陈定国和王一先等同志表示感谢！并特向董火明、胡勇、赵晶、曹薇、赵琴、王晓嘉、谢昭、吴田富、庞博同志表示感谢！他们为本书做出了大量卓有成效的工作，资料的搜集和整理、公式的推导和证明、仿真实例的设计和实现、文档的编排和形成等等无不蕴涵着他们的智慧、浸透着他们的汗水。再次向他们表示衷心感谢！

高 隽

目 录

出版说明

前言

第 1 章	引言	1
1.1	人工神经网络简介	1
1.2	人工神经网络发展历史	1
1.2.1	兴起阶段	1
1.2.2	萧条阶段	3
1.2.3	兴盛阶段	4
1.3	人工神经网络模型	6
1.3.1	生物神经元模型	6
1.3.2	人工神经网络的模型	7
1.4	人工神经网络的分类及学习规则	10
1.4.1	人工神经网络的分类	10
1.4.2	神经网络的学习	12
1.5	人工神经网络的信息处理能力	16
1.6	人工神经网络的应用	17
1.7	人工神经网络与人工智能	18
1.7.1	人工智能简介	18
1.7.2	人工智能与人工神经网络	19
1.8	习题	20
第 2 章	单层前向网络及 LMS 学习算法	23
2.1	单层感知器	23
2.1.1	单层感知器模型	23
2.1.2	单层感知器的学习算法	24
2.2	自适应线性元件	27
2.3	LMS 学习算法	30
2.4	仿真实例	33
2.5	习题	43
第 3 章	多层前向网络及 BP 学习算法	44
3.1	多层感知器	44
3.2	BP 学习算法	45
3.2.1	BP 学习算法	45
3.2.2	BP 学习算法步骤	49
3.2.3	BP 学习算法的改进	53
3.3	径向基网络	55

3.3.1	RBF 神经网络模型	56
3.3.2	RBF 网络的学习算法	59
3.3.3	RBF 网络与多层感知器的比较	63
3.4	仿真实例	63
3.5	习题	73
第 4 章	支持向量机及其学习算法	76
4.1	统计学习理论	77
4.2	支持向量机	78
4.2.1	线性支持向量机	78
4.2.2	非线性支持向量机	81
4.2.3	支持向量机与多层前向网络的比较	84
4.3	支持向量机的学习算法	84
4.3.1	学习算法	84
4.3.2	改进算法	85
4.4	仿真实例	86
4.5	习题	94
第 5 章	Hopfield 神经网络与联想记忆	95
5.1	神经动力学	95
5.2	离散 Hopfield 神经网络	97
5.2.1	离散 Hopfield 网络模型	97
5.2.2	离散 Hopfield 网络的运行规则	98
5.3	连续 Hopfield 神经网络	99
5.3.1	连续 Hopfield 网络模型	100
5.3.2	连续 Hopfield 网络稳定性分析	101
5.4	联想记忆	102
5.4.1	联想记忆的基本概念	102
5.4.2	Hopfield 联想记忆网络	105
5.4.3	Hopfield 联想记忆网络的运行步骤	106
5.4.4	联想记忆网络的改进	108
5.5	仿真实例	109
5.6	习题	118
第 6 章	随机神经网络及模拟退火算法	119
6.1	Boltzmann 机	119
6.1.1	Boltzmann 机的网络结构	120
6.1.2	Boltzmann 机的工作原理	121
6.1.3	Boltzmann 机的运行步骤	123
6.1.4	Boltzmann 机的学习规则	124
6.2	Boltzmann 机的改进	126
6.2.1	确定性 Boltzmann 机	126

6.2.2	Sigmoid 置信度网络	127
6.3	模拟退火算法	129
6.3.1	模拟退火原理	130
6.3.2	模拟退火算法用于组合优化问题	131
6.4	仿真实例	133
6.5	习题	140
第 7 章	竞争神经网络	141
7.1	Hamming 网络	142
7.1.1	Hamming 网的网络结构	142
7.1.2	网络的运行过程	143
7.1.3	网络的学习算法	144
7.2	自组织映射网络	146
7.2.1	自组织映射网络模型	147
7.2.2	自组织映射学习算法	148
7.3	学习矢量量化	151
7.3.1	网络模型	151
7.3.2	学习矢量量化的学习算法	152
7.3.3	学习矢量量化和自组织映射	154
7.4	主分量分析	155
7.4.1	主分量分析方法	155
7.4.2	前向主分量分析网络及其算法	158
7.4.3	自适应主分量网络及其算法	160
7.5	仿真实例	162
7.6	习题	176
第 8 章	协同神经网络	178
8.1	协同学简介	179
8.1.1	协同学的基本概念	180
8.1.2	协同学的数学模型	180
8.2	协同神经网络及其学习算法	183
8.2.1	协同神经网络的数学模型	183
8.2.2	协同神经网络的结构	185
8.2.3	协同神经网络的运行步骤	188
8.2.4	协同学习算法	189
8.3	仿真实例	192
8.4	习题	196
附录 1	MATLAB 及神经网络工具箱简介	197
附录 2	MATLAB (5.X 版) 中神经网络工具箱函数	203
参考文献		208

第1章 引言

1.1 人工神经网络简介

人工神经网络 (Artificial Neural Network, ANN), 亦称为神经网络 (Neural Networks, NN), 是由大量处理单元 (神经元 Neurons) 广泛互连而成的网络, 是对人脑的抽象、简化和模拟, 反映人脑的基本特性。人工神经网络的研究是从人脑的生理结构出发来研究人的智能行为, 模拟人脑信息处理的功能。它是根植于神经科学、数学、统计学、物理学、计算机科学及工程等学科的一种技术。

人工神经网络是由简单的处理单元所组成的大量并行分布的处理机, 这种处理机具有存储和应用经验知识的自然特性, 它与人脑的相似之处概括为两个方面: 一是通过学习过程利用神经网络从外部环境中获取知识; 二是内部神经元 (突触权值) 用来存储获取的知识信息。

神经网络也经常被称为神经计算机 (Neurocomputer), 但它与现代数字计算机的不同之处主要表现在以下方面:

- 1) 神经网络的信息存储与处理 (计算) 是合二为一的, 即信息的存储体现在神经元互连的分布上; 传统计算机的存储与计算是独立的, 因而在存储与计算之间存在着瓶颈。

- 2) 神经网络以大规模模拟计算为主; 数字计算机是以串行离散符号处理为主。

- 3) 神经网络具有很强的鲁棒性和容错性, 善于联想、概括、类比和推广, 任何局部的损伤不会影响整体结果。

- 4) 神经网络具有很强的自学习能力, 能为新的输入产生合理的输出, 可在学习过程中不断完善自己, 具有创新特点。

- 5) 神经网络是一大规模自适应非线性动力系统, 具有集体运算的能力。这与本质上是线性系统的现代数字计算机迥然不同。

人工神经网络是近年来的热点研究领域, 涉及到电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、控制科学与技术等诸多学科, 其应用领域包括: 建模、时间序列分析、模式识别和控制等, 并在不断的拓展。

1.2 人工神经网络发展历史

该项研究始于 20 世纪 40 年代, 它的发展经历了一条由兴起、萧条和兴盛三个阶段构成的曲折道路。

1.2.1 兴起阶段

1943 年, 精神病学家和神经解剖学家 McCulloch 与数学家 Pitts 在数学生物物理学会刊《Bulletin of Mathematical Biophysics》上发表文章。总结了生物神经元的一些基本生理特征,

提出了形式神经元的数学描述与结构,即 MP 模型。McCulloch 和 Pitts 描述了一个逻辑微积分的神经网络,这个网络由神经生理学和数学逻辑学习组成,他们定义的神元元的正规模型认为遵循着“全或无”法则。McCulloch 和 Pitts 表明:原则上,拥有了数量众多的简单单元和适当的神经元联接且运行同步的情况下,所构建的网络能计算任何可计算的函数。MP 的提出兴起了 NN 研究,同时产生了人工智能(Artificial Intelligence, AI)。

1948 年, Wiener 所著的著名的《Cybernetics》一书出版, Wiener 提出了控制、通信和统计信号处理的重要概念,在学科之间抓住了统计方法的物理意义。

1949 年,生理学家 D. O. Hebb 出版了《The Organization of Behavior》一书。该书第一次鲜明提出了改变神经元连接强度的 Hebb 规则,特别是,Hebb 提出脑中互连信息随着感官学习任务的不同而不断变化,这种变化产生了神经集合。他认为学习过程是在突触(Synapse)上发生的,突触的联系强度随其前后神经元的活动而变化。根据这一假设提出的学习规则为神经网络的学习算法奠定了基础,使神经网络的研究进入了一个重要的发展阶段。

1952 年, Ashby 所著的《Design for a Brain: The Orgin od Adaptive Behavior》一书出版。书中论述了自适应行为不是天生的,而是学习的结果这一基本概念,而且通过对动物行为(系统)的学习会有更好的改变,书中还强调了类似机器的生物动态和相关的稳定性概念。

1954 年, Minsky 在 Princeton 大学撰写了一篇题为《Theory od Neural-Analog Reinforcement Systems and Its Application to the Brain-Model Problem》的神经网络博士论文。1961 年, Minsky 写了一篇关于早期人工智能的优秀论文,题为《Steps Toward Artificial Intellience》,论文的后半部包含了当今神经网络的大部分内容;1967 年,出版了 Minsky 的《Computation: Finite and Infinite Machines》一书,该书清晰地扩展了 McCulloch 和 Pitts1943 年的成果,并将其归入自动机和计算理论中。

1954 年,通信理论的先驱和手写器的发明者 Gabor 提出了非线性自适应滤波思想。他希望与合作者一起发明一种机器,通过将一个随机过程的样本输入机器中,连同目标函数,实现机器学习。

Von Neumann 是 20 世纪前半叶科学领域中最伟大的人物之一,数字计算机的基础设计就是以他的名字命名的。早在 1949 年, Von Neumann 在 Illinois 大学四次讲座的第二次讲座中,就阐述了 McCulloch-Pitts 正式的神经网络理论特点。1955 年,他应邀去 Yale 大学进行 Silliman 讲座,直至 1956 年(他死于 1957 年)。他未完成的 Silliman 讲座的手稿于 1958 年作为一本书被出版,书名为《The Computer and the Brain》,此书由于涉及了 Von Neumann 生前所做的工作和他注意到的人脑与计算机的巨大差异,而倍受关注。此外,1956 年 Von Neumann 用约简的思想解决了一个在神经网络中特别令人关注的问题,这就是如何用认为是不可靠的神经元来设计一个可靠的网络问题。

1957 年, Rosenblantt 提出感知机(Perceptron)。1958 年 Rosenblatt 基于对感知器的研究,提出了解决模式识别问题的新的监督学习方法,并证明了所谓的感知器收敛定理。首次把神经网络的研究付诸工程实践。这是一种学习和自组织的心理学模型,它基本上符合神经生理学的知识,模型的学习环境是有噪声的,网络构造中存在随机连接,这是符合动物学习的自然环境。这种类型的机器显然有可能应用于模式识别(Pattern Recognition)、联想记忆(Associative Memory)等方面。

1960 年, Widrow 和 Hoff 引入了最小均方差 (Least Mean-Square, LMS) 算法, 并用它系统阐明了自适应线性元件(Adaptive Linear Element)。感知器和自适应线性元件之间的差异在于训练过程, 最早出现的包含多适应元件的可训练分层神经网络是多学习机结构, 这种结构由 Widrow 和他的学生在 1962 年提出。

1965 年, Milsson 所著的《Learning Machines》一书出版, 至今这本书仍然是关于超平面中线性模型的一本佳作。

1967 年, Amari 用推测梯度方式进行自适应模式分类。

在感知器盛行的 20 世纪 60 年代, 人们对神经网络的研究过于乐观, 认为只要将这种 NN 互连成一个网络, 就可以解决人脑思维的模拟问题。因此, 当时有上百家实验室纷纷投入这项研究, 美国军方也投入了巨额资金, 当时 NN 在声纳信号识别等领域的应用取得了一定的成绩。

但是 1969 年, Minsky 和 Papert 所著的《Perceptron》一书出版, 该书从数学角度证明了关于单层感知器的计算具有根本的局限性, 指出感知机的处理能力有限, 甚至连 XOR 这样的问题也不能解决, 并在多层感知器的总结章中, 论述了单层感知器的所有局限性在多层感知器中是不可能被全部克服的。当时 AI 的以功能模拟为目标的另一分支, 出现了转机, 产生了以知识信息处理为基础的知识工程(Knowledge Engineering), 给 AI 从实验室走向实用带来了希望。同时, 微电子技术的进步, 以及传统的 Von Neumann 型数字计算机发展, 使整个学术界陶醉于数字计算机的成功之中, 从而使 NN 的研究进入了萧条时期。

1.2.2 萧条阶段

根据 Cowan 1990 年提出的观点, 有三个原因导致了神经网络研究的十多年滞后。原因之一是技术上的——没有个人计算机和工作空间站进行实验, 如 Gabor 发展了他的非线性学习滤波器, 却花费了额外的六年时间建造了含有类推装置的滤波器; 原因之二一半是心理上的, 即 Minsky 和 Papert 对感知器的悲观结论, 一半是资金上的, 即没有代理商资助。原因之三是神经网络和晶格旋转之间的类推还未成熟, 直到 1975 年才由 Sherrington 和 Kirkpatrick 创建出旋转镜片模型。

这些因素导致了 20 世纪 70 年代对神经网络的研究陷入了低谷, 此时许多研究人员放弃了除心理学和神经学之外的其他领域。

难能可贵的是, 在此期间, 仍有不少学者在极端艰难的条件下, 保持对神经网络的信奉, 致力于这一研究。

1972 年 Teuvo Kohonen 和 James Anderson 各自独立发展了用于记忆的新的神经网络。Amari 独立提出了一个神经元的附加模型, 并将其应用于研究随机联接类似于神经元元件的动态行为之中。Wilson 和 Cowan 从含有兴奋和抑制神经元空间局部化模型的动态行为中获取非线性微分方程组。

1976 年 Willshaw 和 von der Malsburg 发表了第一篇受人脑拓扑次序映射启发, 构筑自组织映射的论文。

1977 年 Anderson, Silverstein, Ritz 和 Jones 提出了黑箱脑状态 (Brain-State-in-a-Box, BSB) 模型, 其中包括含有非线性动力的一个简单联想网络。

1980 年 Grossberg 在早期对竞争学习研究的基础上, 创立了自组织新理论, 被称为自适

应性谐振理论 (Adaptive Resonance Theory, ART), 理论主要包括自下而上的认知层和自上而下的生成层, 若输入模式与学习反馈模式相匹配, 则产生称为“自适应性谐振”的动力状态 (即神经网络的扩大与延长), 其前向/后向投影原理已经被其他学者应用于不同方面, 使其得到了发展。

由于新思想的缺乏和用于实验的高性能计算机的短缺, 导致了 20 世纪 60 年代后期神经网络研究的停滞不前, 到了 20 世纪 80 年代, 上述难关均被攻克, 神经网络有了迅猛发展, 功能日渐强大的个人计算机和 workstation 开始被广泛的应用, 此时也产生了神经网络研究的新观念。

使神经网络重焕生机的原因, 主要有两个新观念, 第一个是利用统计机制解释循环网络的运行过程, 这种机制可用于一个联想记忆, 此观念是由物理学家 John Hopfield 在他的核心论文中提出的。

20 世纪 80 年代的第二个核心发展是用于训练多层感知器的反向传播算法, 此算法同时被几个学者所发现, 其中影响最大并被广泛传播的是由 David Rumelhart 和 James McClelland 提出的算法, 该算法是对 20 世纪 60 年代 Minsky 和 Papert 的批判的一个有力的回答。

另外, 20 世纪 70 年代后期出现以下情况:

(1) 研究视、听觉的人工智能专家首先遇到挫折, 计算机一般不能从现实世界的实例与现象中获取并总结出知识, 也就是说计算机不具备学习能力。于是人们开始意识到了 Von Neumann 体系结构的局限性, 转而研究数据流机和并行计算机体系结构;

(2) 日本第五代计算机计划远未达到预想水平, 使人们觉得有必要弄清人们习以为常的认知功能是如何进行的, 这些认知功能包括视听觉感知、学习记忆、运行控制等, 从而使人们认识到必须开拓新的思路, 探索新的实现途径——与人脑的生理组织更为接近的 NN 模型自然成为理想的候选模型;

(3) 在人类智能行为研究方面, 神经生理学家、心理学家和计算机科学家相互结合, 认为人脑是一个功能十分强大、结构异常复杂的信息系统, 但其基本仍是神经元及其之间的连接;

(4) NN 的研究依靠众多学科的共同发展, 是多学科的综合产物。而当时多学科都有了相应的发展, 如普里高津 (Prigogine) 提出了非平稳系统的自组织理论, 获诺贝尔奖; 哈肯 (Haken) 研究了大量单元集团运动而产生的宏观效果; 非线性系统“混沌”态的提出及其研究等等, 这些都是研究如何通过单元之间的相互连接作用建立复杂系统, 类似于生物系统的自组织行为;

(5) 脑科学与神经科学的研究成果, 迅速反映到神经网络的改进上, 例如视觉研究中发现的侧抑制原理、感受野的概念, 听觉通道上神经元的自组织排列等, 生物的 NN 研究成果对 ANN 的研究起了重要的推动作用。

所有这些方面重新引起了人们对 ANN 的研究兴趣。

1.2.3 兴盛阶段

学术界公认, 标志 NN 研究高潮的又一次到来是美国加州理工学院生物物理学家 J. Hopfield 教授于 1982 年和 1984 年发表在美国科学院院刊上的两篇文章以及 1986 年 Rumelhart 与 McClelland 的两册书。

1982 年, Hopfield 用能量函数的思想形成了一种新的计算方法, 该计算方法由含有对称突触联接的反馈网络执行。而且他还将该反馈网络同用于统计物理的 Ising 模型相类推, 这种类推为大量的物理学理论和许多的物理学家进入神经网络领域铺平了道路。Hopfield 阐明了 NN 与动力学关系, 并用非线性动力学的方法来研究这种 NN 的特性, 建立了 NN 稳定性判据, 并指出信息存储在网络中 NN 之间的连接上, 形成了 Hopfield 网络。这是 NN 研究的突破性进展。

1982 年, 另一个重要的发展是 Kohonen 的关于自组织图的论文发表, 其中用到了一个一维或二维的晶体结构, 这种模型已成为衡量在此领域中有价值创新的基准。

1983 年, Kirkpatrick, Gelatt 和 Vecchi 提出了模拟退火(Simulated Annealing, SA)的新方法, 该方法以统计理论为基础, 用于解决组合最优问题。

1983 年, Barto, Sutton 和 Anderson 发表了关于强化学习(Reinforcement Learning, RL)的论文, 虽然他们不是第一个提出强化学习思想的人(如 Minsky 在 1954 年的博士论文中有提及), 但这篇论文将强化学习应用于实际, 验证了其应用具有可行性, 对强化学习的发展及应用产生了重要的影响。

1984 年, Hopfield 设计与研制了他所提出的神经网络模型的电路, 指出神经元可以用运算放大器来实现, 所有神经元的连接可用电子线路来模拟。同时, 他也进行了神经网络应用研究, 成功的解决了旅行商(TSP)计算难题(优化问题), 令人震惊, 使神经网络的研究步入了兴盛时期。

1984 年, Braitenberg 出版了《Vehicles: Experiments in Synthetic Psychology》一书, 此书中引用面向对象技术和自组织原理, 通过对假定的基本机制合成, 而不是自顶向下的分析, 使得对复杂过程的理解达到最好效果。Braitenberg 通过对各种具有简单内部结构的机器的描述证明了这一重要法则。Braitenberg 直接和间接的研究动物大脑超过 20 年, 在这方面的研究启发其发现机器特性及其行为。

1985 年, Ackley, Hinton 和 Sejnowski 以模拟退火思想为基础, 对 Hopfield 模型引入了随机机制, 提出了 Boltzmann 机。第一次成功实现了多层神经网络的功能, 打破了人们心理上的局限, 证明了 Minsky 和 Papert 1969 年的推测根据是不正确的。

1986 年, Rumelhart, Hinton 和 Williams 发展了反向传播算法(Back-Propagation algorithm, BP)。同年, Rumelhart 和 McClelland 编写的名为《Parallel Distributed Processing: Explorations in the Microstructures of Cognition》一册两本出版, 此书的出版对反向传播算法的应用产生了重要影响。反向传播算法已成为大多数多层感知器训练所采用的流行学习算法。该算法解决了多层 NN 的学习问题, 证明了多层神经网络的计算能力并不象 Minsky 等人所预料的那样弱, 相反它可以完成许多学习任务, 解决许多实际问题。

1988 年, Linsker 在感知器网络上提出了一种新的自组织理论, 该理论用于保持输入行为模式的最大信息, 并受突触联接和活动范围的限制, 在 Shanno 信息论的基础上, 形成了最大互信息理论。Linsker 的论文重新点燃了基于神经网络的信息应用理论的光芒, 尤其对 Bell 和 Sejnowski 提出的盲源分离问题(Blind Source Separation Problem)的信息应用理论产生了影响, 同时激起许多学者去探索用其他信息理论模型去解决各自广泛领域中的问题, 该方法称为盲反卷积(Blind Deconvolution)。

1988 年, Broomhead 和 Lowe 用径向基函数(Radial Basis Functions, RBF)提出了分层

反馈网络设计的方法。Broomhead 和 Lowe 的论文将神经网络的设计与数值分析的重要领域和线性适应滤波挂钩, 产生了大量的研究成果, 为多层感知器的研究提供又一方法和途径。1990 年, Poggio 和 Girosi 应用 Tikhonov 规则理论进一步丰富了径向基理论。

1989 年, Mead 的《Analog VLSI and Neural System》一书出版, 这本书将生物神经和集成电路结合在一起, 给出了硅视网膜和硅耳蜗。

此后, 神经网络领域的研究有了新发展, 在 20 世纪的最后 10 年中, 产生了大量关于神经网络的论文, 并在许多领域应用了神经网络技术, 新的理论和实践工作层出不穷, 20 世纪 90 年代初期, Vapnik 和合作者们发明了一类计算功能强大的有导师学习网络, 用于解决模式识别、回归及密度估测问题, 这种网络被称为支持向量机(Support Vector Machines, SVM), 以有限样本学习理论的结论为基础。支持向量机的新特征在于 Vapnik-Chervonenkis (VC) 维特征蕴含在向量机的设计中, VC 维数为衡量神经网络样本学习能力提供了一种有效的量度。

总之, 以 Hopfield 教授 1982 年发表的论文为标志, 掀起了神经网络的研究热潮。1987 年 6 月, 在美国加州举行了第一届 NN 国际会议, 有一千多名学者参加, 并成立了国际 NN 学会, 以后每年召开两次国际联合 NN 大会 (IJCNN), 其他国际学术会议也都列有 NN 主题。较有影响的国际学术刊物有:《IEEE Transaction on Neural Network》和《Neural Network》。美国 IBM、AT&T、贝尔实验室、神经计算机公司、各高校、美国政府制定了“神经、信息、行为科学 (NIBS)”计划, 投资 5.5 亿美元作为第六代计算机的基础研究; 美国科学基金会 (NSF)、海军研究局 (ONR) 和空军科学研究部 (AFOSR) 三家投资一千万美元; 美国国防部 DARPA 认为 NN “看来是解决机器智能的惟一希望” “这是一项比原子弹工程更重要的技术” 投资四亿美元。主要研究目标: 目标识别与跟踪, 连续语音识别, 声纳信号辨别。

日本的富士通、日本电气、日立、三菱、东芝急起直追。1988 年日本提出了所谓的人类尖端科学计划 (Human Frontier Science Program), 即第六代计算机研究计划。法国提出了“尤里卡”计划, 还有德国的“欧洲防御”和前苏联的“高技术发展”等等。

我国于 1989 年在北京召开了一个非正式的 NN 会议; 1990 年 12 月在北京召开了中国 NN 大会; 1991 年在南京成立了中国 NN 学会, 由国内十五个一级学会共同发起“携手探智能, 联盟攻大关”的 863 高技术研究计划; 自然科学基金、国防科技预研基金也都列入了 NN 研究内容。

1.3 人工神经网络模型

1.3.1 生物神经元模型

在人类大脑皮层中大约有 100 亿个神经元, 60 万亿个神经突触以及它们的联接体。单个神经元处理一个事件需要 10^{-3} 秒, 而在硅芯片中处理一事件只需 10^{-9} 秒。但人脑是一个非常高效的结构, 大脑中每秒钟每个动作的能量约为 10^{-16} 焦耳, 而当今性能最好的计算机进行相应的操作需要 10^{-6} 焦耳。

神经元是基本的信息处理单元。生物神经元主要由树突、轴突和突触组成。其结构示意图如图 1-1 所示。

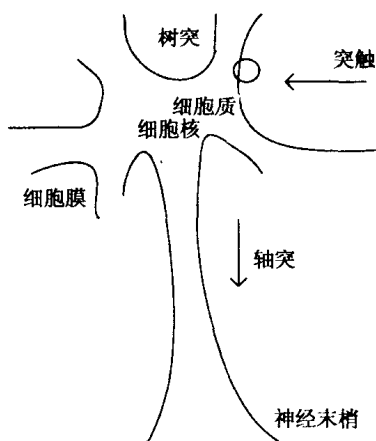


图 1-1 生物神经元模型

其中树突是由细胞体向外伸出的，有不规则的表面和许多较短的分支。树突相当于信号的输入端，用于接受神经冲动。轴突由细胞体向外伸出的最长的一条分支，即神经纤维。相当于信号的输出电缆，其端部的许多神经末梢为信号输出端子，用于传出神经冲动。神经元之间通过轴突（输出）和树突（输入）相互联结，其接口称为突触。每个细胞约有 $10^3 \sim 10^4$ 个突触。神经突触是调整神经元之间相互作用的基本结构和功能单元，最通常的一种神经突触是化学神经突触，它将得到的电信号转化成化学信号，再将化学信号转化成电信号输出。这相当于双接口设备。它能加强兴奋或抑制作用，但两者不能同时发生。细胞膜内外有电位差，约为 $20 \sim 100\text{mv}$ ，称为膜电位。膜外为正，膜内为负。

神经元作为信息处理的基本单元，具有如下重要的功能。

1) 可塑性：可塑性反映在新突触的产生和现有神经突触的调整上，可塑性使神经网络能够适应周围的环境。

2) 时空整合功能：时间整合功能表现在不同时间、同一突触上；空间整合功能表现在同一时间、不同突触上。

3) 兴奋与抑制状态：当传入冲动的时空整合结果，使细胞膜电位升高，超过被称为动作电位的阈值（约为 40mv ），细胞进入兴奋状态，产生神经冲动，由轴突输出；同样，当膜电位低于阈值时，无神经冲动输出，细胞进入抑制状态。

4) 脉冲与电位转换：沿神经纤维传递的电脉冲为等幅、恒宽、编码（ $60 \sim 100\text{mv}$ ）的离散脉冲信号，而细胞电位变化为连续信号。在突触接口处进行“数/模”转换。神经元中的轴突非常长和窄，具有电阻高、电压大的特性，因此轴突可以建模成阻容传播电路。

5) 突触的延时和不应用期：突触对神经冲动的传递具有延时和不应用期，在相邻的二次冲动之间需要一个时间间隔。在此期间对激励不响应，不能传递神经冲动。

6) 学习、遗忘和疲劳：突触的传递作用有学习、遗忘和疲劳过程。

1.3.2 人工神经网络的模型

人工神经网络是由大量处理单元广泛互连而成的网络，是人脑的抽象、简化、模拟，反映人脑的基本特性。一般来说，作为神经元模型应具备三个要素：

(1) 具有一组突触或联接，常用 w_{ij} 表示神经元 i 和神经元 j 之间的联接强度，或称之

为权值。与人脑神经元不同，人工神经元权值的取值可在负值与正值之间。

(2) 具有反映生物神经元时空整合功能的输入信号累加器。

(3) 具有一个激励函数用于限制神经元输出。激励函数将输出信号压缩（限制）在一个允许范围内，使其成为有限值，通常，神经元输出的扩充范围在 $[0,1]$ 或 $[-1,1]$ 闭区间。

一个典型的人工神经元模型如图 1-2 所示。

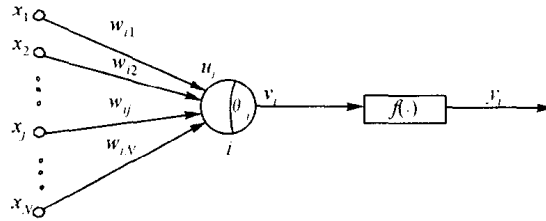


图 1-2 人工神经元模型

其中 x_j ($j=1,2,\dots,N$) 为神经元 i 的输入信号， w_{ij} 为突触强度或联接权。 u_i 是由输入信号线性组合后的输出，是神经元 i 的净输入。 θ_i 为神经元的阈值或称为偏差用 b_i 表示， v_i 为经偏差调整后的值，也称为神经元的局部感应区。

$$u_i = \sum_j w_{ij} x_j \quad (1-1)$$

$$v_i = u_i + b_i \quad (1-2)$$

$f(\cdot)$ 是激励函数， y_i 是神经元 i 的输出。

$$y_i = f\left(\sum_j w_{ij} x_j + b_i\right) \quad (1-3)$$

激励函数 $f(\cdot)$ 可取不同的函数，但常用的基本激励函数有以下三种：

(1) 阈值函数 (Threshold Function)

$$f(v) = \begin{cases} 1, & \text{若 } v \geq 0 \\ 0, & \text{若 } v < 0 \end{cases} \quad (1-4)$$

该函数通常也称为阶跃函数，常用 $u(t)$ 表示，如图 1-3 a 所示。若激励函数采用阶跃函数，则图 1-2 所示的人工神经元模型即为著名的 MP(McCulloch-Pitts)模型。此时神经元的输出取 1 或 0，反映了神经元的兴奋或抑制。

此外，符号函数 $\text{Sgn}(t)$ 也常常作为神经元的激励函数，如图 1-3 b 所示。

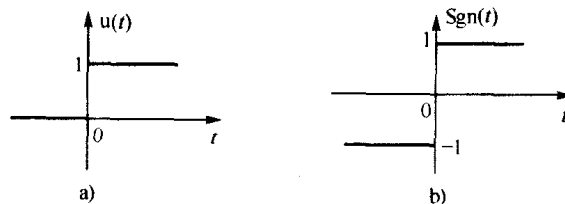


图 1-3 阈值函数