

RENGONG SHENJING WANGLUO JICHU



RENGONG SHENJING WANGLUO JICHU



图书在版编目(CIP)数据

神经网络基础 / 韩敏编著. — 大连 : 大连理工大学出版社, 2014. 3

ISBN 978-7-5611-8933-7

I. ①人… II. ①韩… III. ①神经网络—高等学校—教材 IV. ①TP183

中国版本图书馆 CIP 数据核字(2014)第 043411 号



大连理工大学出版社出版

地址:大连市软件园路 80 号 邮政编码:116023

发行:0411-84708842 邮购:0411-84703636 传真:0411-84701466

E-mail:dutp@dutp.cn URL:http://www.dutp.cn

大连美跃彩色印刷有限公司印刷 大连理工大学出版社发行

幅面尺寸:185mm×260mm 印张:9.75 字数:224 千字

印数:1~1000

2014 年 3 月第 1 版

2014 年 3 月第 1 次印刷

责任编辑:唐 爽

责任校对:张 檬

封面设计:张 莹

ISBN 978-7-5611-8933-7

定 价:25.00 元

前言

多年来,人们从医学、生物学、信息学、计算机科学、认知学、组织协同学等各个角度试图对人体最高等级的器官——“大脑”及其辅助组织神经系统进行探索与模仿。在这样的背景下,诞生了一门新兴的学科——“人工神经网络”。人工神经网络(Artificial Neural Network, ANN),简称神经网络(Neural Network, NN),它是由大量处理单元互联组成的非线性、自适应信息处理系统。人工神经网络是以大脑的工作模式为基础来研究自适应、非程序的信息处理方法。它具有强大的非线性建模能力、自学习能力和容错能力等诸多显著的优点,被广泛地应用于科研以及工程中,并在模式识别与图像处理、控制与优化、金融预测与管理以及数据挖掘等领域取得了良好的效果。

本书作为人工神经网络课程的教材,立足于基础。全书理论完整、层次清晰、内容由浅入深,循序渐进地对人工神经网络的相关知识进行讲解。本书在写作过程中非常注重内容的选择和编排,围绕从理论到实践的主线,从介绍人工神经网络的背景和发展历史入手,首先讲解人工神经网络的特点和功能,并强调其在诸多领域的应用,进而引起读者深入研究的兴趣。之后辅助以生物神经网络,来加强读者对人工神经网络一些基本概念的理解。以网络结构、学习算法、工作原理为核心详细介绍典型人工神经网络模型,同时加入人工神经网络的应用实例,使读者在理论与应用层面均有所收获。

作者长期致力于人工神经网络及其学习算法的研究,本书正是结合了作者多年的研究和教学实践经验,并参考国内外学者编著的相关书籍和文献资料最终编写完成的。作者力图为高等院校工科专业的研究生和本科生,以及相关专业领域的科研人员和工程技术人员提供一本介绍人工神经网络基本理论和典型模型的系统性教材和参考书籍。

全书共分为七章。第一章神经网络概述,从神经网络的基本概念、起源、发展的几度起伏以及基本特点出发,对神经网络进行全面且深入的介绍。第二章神经网络基础,

主要阐述了人工神经网络的基础知识,从生物学角度出发,介绍了包括生物神经元、人工神经元、人工神经网络模型及其常用的学习算法与泛化能力。第三章前馈神经网络、第四章递归神经网络、第五章自组织竞争型神经网络和第六章支持向量机,分别就人工神经网络中十一个重要的网络模型进行了深入的探讨,并对每种模型的基本概念、结构、算法以及实现步骤进行讲解。全书最后一章中辅以应用实例,以方便读者进行仿真实现并加深对神经网络的理解。

在本书出版之际,作者要感谢参与本书相关内容的研究及书稿编写、校对工作的老师和学生们,他们是秦攀副教授,博士生张檬、许美玲,硕士生任伟杰、曹占吉、边茂松、张雅美、张成坤、徐旭东。全书由博士生张檬做了认真细致的统稿。另外,也要感谢大连理工大学出版社的赵静老师对本书编写工作的帮助与支持。

由于作者水平有限,书中难免存在错误和不足之处,敬请广大读者批评指正,以便进一步修订完善。

韩 敏

2014年1月21日于大连



第1章 神经网络概述	1
1.1 神经网络的基本概念	1
1.1.1 人脑与电子计算机的比较	2
1.1.2 什么是神经网络	3
1.2 神经网络发展的历史回顾	4
1.2.1 神经网络研究的兴起	4
1.2.2 神经网络研究的萧条与反思	6
1.2.3 神经网络研究的复兴与再发展	8
1.2.4 神经网络研究的热潮	9
1.2.5 神经网络的新发展阶段	10
1.3 神经网络的研究进展	12
1.4 神经网络的基本特点、功能与应用	14
1.4.1 神经网络的基本特点	14
1.4.2 神经网络的基本功能	16
1.4.3 神经网络的应用	17
本章小结	19
思考与练习	20
第2章 神经网络基础	21
2.1 人脑神经系统概述	21
2.1.1 人脑的基本结构	21
2.1.2 人脑神经系统的各部分功能	23
2.2 生物神经网络的基本概念	26
2.2.1 生物神经元	26
2.2.2 生物神经元的信息传递方式	28
2.3 人工神经网络的基本概念	29
2.3.1 人工神经元	29
2.3.2 神经元的激活函数	30
2.4 人工神经网络模型	32
2.4.1 神经网络的网络结构	32
2.4.2 典型的神经网络结构模型	34
2.5 神经网络的学习	35
2.5.1 神经网络学习规则分类	36
2.5.2 几种常用的学习规则	37

2.5.3 神经网络的泛化能力	41
本章小结	42
思考与练习	43
第3章 前馈神经网络	44
3.1 基本概念	44
3.2 感知器	45
3.2.1 单层感知器	45
3.2.2 多层感知器	48
3.3 BP神经网络	49
3.3.1 误差反向传播算法	49
3.3.2 误差反向传播算法的改进	53
3.3.3 BP神经网络的设计	57
3.4 径向基函数神经网络	59
3.4.1 径向基函数神经网络的结构	59
3.4.2 径向基函数神经网络的学习算法	61
本章小结	63
思考与练习	63
第4章 递归神经网络	65
4.1 递归神经网络的基本概念	65
4.2 Hopfield神经网络	66
4.2.1 离散型 Hopfield神经网络	66
4.2.2 连续型 Hopfield神经网络	69
4.3 Boltzmann机	72
4.3.1 Boltzmann机的基本概念	72
4.3.2 Boltzmann机网络结构	73
4.3.3 Boltzmann机学习算法	76
4.4 储备池网络	77
4.4.1 储备池网络结构模型	77
4.4.2 储备池网络学习算法	82
本章小结	84
思考与练习	84
第5章 自组织竞争型神经网络	86
5.1 自组织竞争型神经网络基本概念	86
5.1.1 内星和外星学习规则	87
5.1.2 竞争学习规则	87
5.2 自适应共振理论神经网络	88
5.2.1 ARTI型神经网络的结构及工作原理	89
5.2.2 ARTI型神经网络的学习规则	91

5.3 自组织特征映射神经网络	93
5.3.1 自组织特征映射神经网络模型	93
5.3.2 自组织特征映射神经网络运行原理	95
5.3.3 自组织特征映射神经网络学习算法	95
5.4 对向传播神经网络	97
5.4.1 对向传播神经网络结构	97
5.4.2 对向传播神经网络训练过程	97
5.5 学习向量量化神经网络	98
5.5.1 向量量化	99
5.5.2 学习向量量化神经网络工作原理	99
本章小结	100
思考与练习	101
第6章 支持向量机	102
6.1 支持向量机的基本概念	102
6.2 线性支持向量机	103
6.2.1 线性可分离问题的算法	103
6.2.2 线性不可分离问题的算法	110
6.3 非线性支持向量机	112
6.3.1 非线性问题的算法	112
6.3.2 核函数	114
本章小结	115
思考与练习	116
第7章 神经网络仿真实例	117
7.1 仿真数据说明	117
7.2 BP神经网络仿真实例	121
7.3 径向基函数神经网络仿真实例	123
7.4 Hopfield神经网络仿真实例	125
7.5 Boltzmann机仿真实例	126
7.6 储备池网络仿真实例	127
7.7 ARTI型神经网络仿真实例	129
7.8 自组织特征映射神经网络仿真实例	133
7.9 对向传播神经网络仿真实例	134
7.10 学习向量量化神经网络仿真实例	137
7.11 支持向量机仿真实例	139
本章小结	141
思考与练习	142
附 录	143
参考文献	145

第 1 章

神经网络概述

早在古希腊时期,人们就已经开始了针对认知和思维机器的探索研究。1946 年第一台电子计算机的诞生标志着在此方面的研究有了突破性的进展。随着研究的逐步深入,人们深刻地意识到思维机器、智能计算机的研究对人类社会的发展具有的重要意义。计算机单纯的计算功能已无法满足实际的需求,人们希望它能够具有更加智能的特性,能解决更为复杂且具有逻辑性的一些问题。针对以上需求,在 1956 年,人工智能(Artificial Intelligence, AI)的概念被引入,它主要是研究利用计算机来模仿大脑进行推理、设计、思考、学习等思维活动,进而解决和处理实际生活中一些较为复杂的问题。人工智能的研究目的主要包括以下两个方面:一是提高人类认识自然、改造自然以及推动社会发展的能力;二是让人类进一步认识自己。其研究方法包括实现功能的模拟和实现生理结构的模拟两种。

神经网络可以分为生物神经网络(Biological Neural Network, BNN)和人工神经网络(Artificial Neural Network, ANN)。生物神经网络是指存在于大脑中的实际神经网络;而仿照生物神经网络构造出来的网络为人工神经网络。在工程与学术界中所提到的神经网络一般是指人工神经网络,常简称其为神经网络。神经网络是一种典型的通过生理结构模拟实现人工智能的方法,它是从人脑的生理结构出发,利用仿生学的观点,探索人类智能活动的机理,将对人脑微观结构的研究与其智能行为的研究相结合的一种方法。神经网络系统具有高度的非线性及自适应自组织的特性,它常被用来模拟认知、决策和控制等智能行为。神经网络理论的研究为人们解决大规模信息的并行处理和并行计算问题奠定了基础。自 20 世纪 80 年代以来,神经网络理论进入了一个快速发展的时期,引起了众多领域学者的关注,渗透到了几乎所有的工程应用领域。迄今,神经网络已然成为世界范围内迅速发展起来的一个前沿研究课题,具有相当繁杂庞大的理论和技术内容,取得了诸多具有重大意义的研究成果。

对于初学神经网络理论的读者,笔者建议应首先掌握最基本的概念,包括究竟什么是神经网络?神经网络理论是在怎样的科学背景下产生的?它的发展历史、特点以及基本结构是怎样的?对它的研究究竟有什么意义?对于这些问题,本章将一一进行介绍。

1.1 神经网络的基本概念

电子计算机发展至今,其在计算速度或者按照某种编程规则执行复杂运算方面已经

可以做到非常精确与高效,在模拟人脑逻辑思维方面也有了长足的进展。对于一些较复杂的计算问题,电子计算机是可以完全战胜人类的。但是一旦涉及形象思维和灵感思维一些需要联想或经验判断的问题,电子计算机就远不及人脑的反应速度。例如通过计算机来识别人脸、智能机器如何接住抛出的球等问题。为什么这些人类可以凭直觉与经验很快做出判断的动作,对于电子计算机来说却是异常困难呢?如果将人脑与电子计算机的工作方式进行对比,就不难找到答案。

1.1.1 人脑与电子计算机的比较

人类由于拥有高度发达的大脑而优越于其他的生物。人脑是一个极其庞大、复杂并且功能完善的系统,它是人类一切思维活动的基础,而思维则是人类智能的集中体现。人脑至少包含逻辑和形象两种基本思维方式。其中,逻辑思维的基础主要是概念、规则以及推理,即先将信息抽象为概念,再按照逻辑规则进行推理。由于概念可以用符号表示,而逻辑推理是按串行方式进行的,那么逻辑思维的过程可以事先写成串行指令然后交由机器完成,这正是电子计算机模拟人脑逻辑思维的方式。但是很明显计算机只能解决那些特征明确、推理或运算规则清楚的可编程问题,而对于一些与形象思维和灵感思维相关的问题,例如人脸识别以及运动等涉及联想或通过经验判断的问题,计算机就显得毫无天赋,很难完成。

人们将人脑的计算方式与传统的数字计算机相对比,从而发现问题根源所在。在对比过程中,人们发现无论是在信息处理的能力方面,还是在信息处理的机制方面两者都存在很大的差异。虽然现代的电子计算机发展迅速,针对数值运算、推理及逻辑判断等复杂的问题均可以很迅速的完成,但是与人脑相比,还是存在很大的差距。

1. 对于信息的处理,人脑不仅能对已学习的知识进行记忆,而且可以通过对外界输入信息的判断,联想到一系列早已存储在脑中的相关信息,从而将思维信息串联。这是一个积累的过程,是进行联想和创造的基础。对于电子计算机,人们输入的信息是保持不变的,因此不会出现信息遗忘的问题。如果我们在同样的地址存入新的信息,那么之前的信息将被覆盖,这决定了它没有回忆和积累的可能。另外,计算机内部各存储单元之间是没有关联性的,因此也没有联想的能力,没有积累和联想能力,也就决定了它无法进行创造的本质。

2. 人脑具有从实践中不断学习新的知识和技能、总结经验和教训的能力,这是一个认识、改变和成长的过程,体现的是人脑的学习与认知的能力。而计算机的工作过程都是人们预先设计好的,它按照编写的程序进行,最终得到预先设定好的结果,这只是一个执行的过程,并不会进行学习和积累经验。

3. 对信息进行加工和处理的综合能力方面,人脑善于对客观世界的各种信息和知识进行归纳、对比和总结。人脑具有复杂的非逻辑的信息加工能力,包括具有联想和想象等需要形象思维和灵感思维才能实现的一些功能。也就是说,人脑的功能并不局限于一些数值和逻辑运算,它在解决复杂的实际问题时,不仅要考虑逻辑思维的原则,还要考虑经验、直觉等非逻辑的原则,进而做出综合性的判断。这种综合的判断能力也是人脑创造能

力的基础。电子计算机没有非逻辑加工和处理的综合能力,它被有限条件下逻辑的认识所限制,缺乏辩证逻辑能力。计算机的信息综合能力取决于它所执行的程序,由于没有经验的积累和形象的思维能力,也就不能达到人脑所具有的融会贯通的信息综合处理能力。

人们普遍认为人脑对信息的处理速度应该远慢于电子计算机。事实上,这个观点是片面的,如果只考虑数值处理等一些只需要串行算法就能解决的问题,那确实如此。然而一旦涉及形象思维、经验与直觉的判断方面,这些人脑很快且很容易就可以解决的问题,计算机却要用其几倍、几十倍的时间来完成。例如,利用计算机进行文字、图像等信息的处理,人脸的识别和接住一个从远处抛来的物体等动作,这些通过人脑的判断做起来很容易的事情,通过计算机处理就变得异常的艰难。

人脑与计算机信息处理能力的差异,特别是形象思维能力的差异主要来源于两者系统结构和信息处理机制的不同,其主要表现可以总结为以下几点:①在系统的结构方面,人脑具有一个规模庞大且构造精细的群体结构,即生物神经网络。生物神经网络是由数百亿神经元相互连接组合而成的,每一个神经元都有一种基本的功能,它可以进行各种极其复杂的思维活动。而计算机是单纯的依据算法程序的存取进行工作的。②在信息的存储方式方面,人脑具有联想的能力,它的信息存储方式是分布式的,可以进行回忆和联想。而显然计算机是不具备这种功能的,它的信息一般集中存储在一个特定区域,并且信息之间是相互独立的。③在信息的表达形式方面,人脑的信号表达形式具有模糊性的特点,可以对模糊的、非逻辑性的问题进行处理,具有信息的综合加工处理能力。计算机中的信息要求必须用确定的逻辑表达式来表示,显然对于非逻辑性的一些问题,它很难进行处理。④在信息的处理机制方面,人脑可以采用信息存储与信息处理一体化的协同并行处理机制。由于人脑具有记忆与联想的能力,新输入的信息要受到原有的存储信息的影响,而处理后的信息留在神经元中成为记忆,利用这种形式可以快速完成各种极其复杂的信息识别和处理任务。相比之下,采用有限集中的串行信息处理机制进行信息处理的电子计算机就显得“笨拙”得多。它是将所有的信息都集中在一个或者几个CPU中进行处理,这种处理机制的时间利用率低,当信息量很大时就会出现速度缓慢等问题。通过上述讨论,不难发现计算机和人脑所采取的信息处理机制的不同才是计算机难以解决具有形象思维特点的问题的根本原因。

1.1.2 什么是神经网络

自从认识到人脑的信息处理机制与传统的数字计算机相比是完全不同的方式开始,人们就掀起了神经网络研究的热潮。对于神经网络的概念,可以从人脑与神经网络的相似性方面进行理解。简单地讲,神经网络就是对生物大脑结构和功能的模仿,并采用数学和物理的方法对其进行研究,进而构成一种信息处理系统。

人一出生就拥有精巧构造的大脑,在成长过程中,可以通过对外界环境的感知以及有意识的训练,进而积累“经验”,建立自身规则。经验是通过时间积累的,一个“发展”中的

神经元是与可塑的人脑同义的。神经网络在两个方面与人脑相似:一是神经网络获取知识的途径与人脑相类似,都是从外界环境中学习得来的。二是互连神经元的连接强度,即突触权值,都是用于储存获取的知识。这里用于完成学习过程的程序称为学习算法,其功能是以有序的方式改变网络的突触权值以获得想要的设计目标。最简单的神经网络是针对一些特定任务或功能进行建模的机器。另外一些较复杂的神经网络,则需要通过学习的过程来实现一些复杂操作。为了获得好的结果,神经网络将一些简单的计算单元相互连接,这些简单的计算单元称为“神经元”或者“处理单元”。据此可以总结出神经网络被看作一种自适应机器的定义:神经网络是一个由简单处理单元构成的规模宏大的并行分布式处理器,具有存储经验知识和使之可用的特性。

1988年,美国神经网络学家 H. Nielsen 曾对神经网络给出了一个较为详细的定义:“神经网络是一个并行、分布处理结构,它由处理单元及被称为连接的无向信号通道互连而成。这些处理单元具有局部内存,可以完成局部操作。每个处理单元有一个单一的输出连接,这个输出可以根据需要被分支成期望个数的许多并行连接,且这些并行连接都输出相同的信号,即相应处理单元的信号,信号的大小不因分支的多少而变化。处理单元的输出信号可以是任何需要的数学模型,每个处理单元中进行的操作必须是完全局部的。也就是说,它必须仅仅依赖经过输入连接到达处理单元的所有输入信号的当前值和存储在处理单元局部内存中的值。”这里提到的处理单元就是指神经网络中的神经元,该定义主要强调了四个方面的内容:并行及分布式处理结构;处理单元的输出可被任意分支;输出信号可为任意数学模型;处理单元的操作是完全局部的。人们将 H. Nielsen 的表述简单概括为:神经网络是由许多个简单的处理单元按照某种特定方式相互连接而成的计算系统,该系统依靠其状态对外部输入信息的动态响应完成信息的处理。

最后需要说明的是,神经网络虽然根据自身的一些特性和能力可以解决一些较复杂的问题,然而在实际问题中,神经网络往往不能单独处理一个问题,通常情况下它都是被整合到协调一致的系统工程方法中进行工作的。具体而言,复杂的实际问题往往被分解成若干个相对简单的子问题,而神经网络则负责处理与其能力相符的子问题。

1.2 神经网络发展的历史回顾

神经网络的研究始于19世纪末期,1890年美国著名心理学家 W. James 发表的一部关于人脑结构与功能的研究专著,拉开了神经网络研究的序幕。神经网络的发展道路是十分曲折的,下面对其发展历史进行简单介绍。

1.2.1 神经网络研究的兴起

1890年,专著《心理学原理》(Principles of Psychology)问世,这也是第一部详细论述

人脑结构及功能的研究专著, W. James 对相关学习及联想记忆的基本原理做了开创性的研究。W. James 在书中表述:“让我们假设所有我们的后继推理的基础遵循这样的规则:当两个基本的脑细胞曾经一起或相继被激活过,其中一个受刺激重新激活时会将刺激传播到另一个。”并且他还指出,大脑皮层上任意一点的刺激量是其他所有发射点进入该点的刺激量的总和。

1943年,在《数学和生物物理学会刊》(Bulletin of Mathematical Biophysics)上美国生理学家 W. S. McCulloch 和数学家 W. Pitts 共同发表了一篇神经网络方面的开创性文章。他们用逻辑数学工具研究客观事件在形成神经网络中的数学模型表达,提出形式神经元的数学模型,即著名的 McCulloch-Pitts 模型,简称为 M-P 模型。该模型除连接权值不被调整外,其他与现在的神经元模型基本相同。尽管现在看来 M-P 模型过于简单,而且其观点也并非完全正确,但它的出现开创了神经网络研究的先河,为以后的研究工作提供了依据。并且 W. S. McCulloch 和 W. Pitts 还证明了任意有限的逻辑表达式都可以由 M-P 模型组成的人工神经网络来实现;另外,他们也是继 W. James 之后最早使用大规模并行计算结构来描述神经元和神经网络的学者。鉴于上述理论贡献, M-P 模型被认为开创了神经网络理论研究的新时代。

1949年,加拿大心理学家、神经生物学家 D. O. Hebb 出版了一本名为《行为组织:神经心理学理论》(The Organization of Behavior: a Neuropsychological Theory)的著作。在该著作中提出了 Hebb 学习规则,为神经网络的学习算法研究奠定了基础。他指出大脑的活动是靠脑细胞的组合连接实现的,对应提出了“连接主义”(Connectionism)这一名词。D. O. Hebb 认为,在神经网络中,信息是存储在连接权值中的。同时他提出了三点假设:①假设连接是对称的,也就是认为从神经元 A 到 B 与 B 到 A 的连接权值是相同的;②假设连接权值的学习速率正比于神经元各激活值之积;③假设学习会对连接权值的强度和类型产生影响,使其发生变化,且这些变化建立起细胞间的连接。D. O. Hebb 的这些观点和假设在后期的研究中都在一定程度上得到了实现。

20世纪50年代初,神经网络理论具备了进行初步模拟实验的条件。美国最早的编程人员之一 N. Rochester、之后被人们称为“遗传算法之父”的 J. Holland 与 IBM 公司的研究人员合作,他们通过网络吸取经验来调节突触强度,并以这种方式模拟 Hebb 的学习规则,开发了一个神经网络模拟程序——“概念者”。之后将该程序在 IBM701 计算机上运行,并取得了成功。在最后的結果中出现了许多突现现象,几乎含有大脑的处理风格。但这个过程中最大规模的模拟神经网络也只有 1 000 个神经元,而每个神经元又只有 16 个结合点,再继续进行试验,便受到计算机的限制。Hebb 的学习规则还影响了当时正在 IBM 实习的研究生 J. McCarthy,他加入了 IBM 的一个小组,探讨有关游戏的智能程序,后来他成为人工智能的主要创始人之一。

1952年,英国生物学家 D. Hodgkin 和 T. H. Huxley 建立了著名的 Hodgkin-Huxley 非线性动力学微分方程,即 H-H 方程,成功地表述了神经放电的电化学机制。该模型用来描述神经膜中所发生的非线性现象,如自激振荡、混沌及多重稳定性问题,为人们探

索神经元的兴奋性提供了基本框架,具有重大的理论意义与应用价值。

1954年,美国犹他大学生物学家 J. C. Eccles 提出了真实突触的分流模型,并通过突触的电生理实验进行了验证,该模型为之后研究神经网络模拟突触的功能提供了原型以及生理学的依据。

1958年,美国计算机学家 F. Rosenblatt 在原始 M-P 模型的基础上引入了学习机制,提出了一种具有学习能力的“感知器”(perceptron)模型,这种模型是具有三层网络特性的神经网络结构。至此,人们完成了从单个神经元到三层神经网络的过渡,得到了第一个完整的人工神经网络模型。这个模型由简单的阈值神经元构成,初步具备了诸如并行处理、分布存储和学习等神经网络的一些基本特性,也被人们称为“学习的机器”,它为从系统角度研究神经网络奠定了基础。F. Rosenblatt 利用感知器模型解决了两个问题:一是信息存储采用怎样的形式;二是如何存储影响认知和行为的信息。通过对这两个问题的回答,F. Rosenblatt 提出这种原始的感知学习机在激励-响应特性方面是“自组织”或“自联合”的。

1960年,美国电机工程师 B. Widrow 和 M. Hoff 发表了一篇对神经网络发展影响极大的文章,《自适应开关电路》(Adaptive Switching Circuits)。文中提出了一种连续取值的线性加权求和阈值网络,即自适应线性元件(Adaptive Linear Element)网络,也可以看成是感知器的变形。它被成功地应用于自适应信号处理和雷达天线控制等连续可调过程。需要强调的是,在研究过程中,创造性地应用了一种后期被人们熟知的 Widrow-Hoff 学习规则,这种规则也称为最小均方差(Least Mean Square, LMS)规则,在数学上就是人们熟知的梯度下降法。B. Widrow 和 M. Hoff 在计算机上设计了仿真的人工神经网络,并在之后利用硬件电路实现了该设计。他们的研究作为今天用大规模集成电路实现神经网络计算奠定了基础,也为神经网络硬件的开发做出了杰出的贡献。

1.2.2 神经网络研究的萧条与反思

1969年,人工智能概念的提出者、美国著名人工智能学者 M. Minsky 和 S. Papert 在深入研究以感知器为代表的神经网络系统的功能及其局限性的基础上,发表了一本对神经网络进行评论的书《感知器》(Perceptrons),该书一经发表即引起了轰动。这本书出版于20世纪60年代末,正是在 B. Widrow 和 M. Hoff 发表了《自适应开关电路》一文之后,此时神经网络仍处于第一次的研究高潮期。在当时美国军方认为神经网络工程应当比“原子弹工程”更重要,这也促使一些科学家怀着极大的热情追求神经网络那遥远但并非不可及的目标。许多研究者抱着对神经网络的学习能力极其乐观的态度,开始对神经网络展开深入的研究,然而却遇到了来自各个方面的困难,使得一些人产生了怀疑和失望。正如 M. Minsky 和 S. Papert 在书中指出的那样,简单的线性感知器的功能有限,并不能解决具有线性不可分特点的二类样本的分类问题,例如简单的线性感知器不可能实现“异或”的逻辑关系。并且对于非线性问题,神经网络应该具有隐含层,但是在理论上还不能

证明将感知机模型扩展到多层网络是有意义的。这些悲观的论点对当时神经网络的研究造成了极大的影响,带来了沉重的打击。雪上加霜的是,同时期在集成电路和微电子技术发展的影响下,传统的冯·诺依曼型(Von Neumann)计算机进入发展的全盛时期,基于符号处理的人工智能迅速发展,这也致使人们将目光和注意力都转移到了人工智能与计算机方面,很多专家学者纷纷放弃了神经网络而转向人工智能等方面课题的研究,这样进一步推动了人工智能的发展,使它处于了主导地位。在之后的十几年里,美国及苏联的一些科研机构纷纷终止了对神经网络相关项目的资助,这使得神经网络进入了漫长的十年萧条期。

在神经网络处于萧条期的这十年中,研究人员骤减,相关研究进展缓慢,这样艰苦的条件更加衬托出那些仍坚持该领域研究的少数科学家的远见卓识,正是他们在此期间的坚持和奉献,为神经网络研究的复兴与新时期的发展奠定了理论基础。

1972年,芬兰学者 T. Kohonen 教授,提出了自组织特征映射神经网络(Self-Organizing Feature Map, SOM),并称该神经网络结构为“联想存储器”(associative memory),这种网络结构一直被沿用至今。对于神经网络的学习方式,可以将其分为有教师(supervised)学习和无教师(unsupervised)学习两种。其中无教师学习是利用网络具有的“记忆”能力对曾学习过的或与其内容相似的刺激进行回忆,最终产生相应的输出。自组织神经网络模型是具有代表性的无教师学习网络,采用了一种“胜者为王”(Winner Take All, WTA)的竞争学习规则,与之前提出的感知器模型有很大的不同,这是一种自组织网络。在现实中,该网络被主要应用于模式识别、语音识别、分类等场合。同年,另一位美国的神经生理学家、心理学家 J. Anderson,提出了一个与自组织神经网络模型非常类似的神经网络,称为“交互存储器”(Interactive Memory),主要针对模型、网络结构、学习规则的生物仿真性等方面做了研究。这里需要提到的是,区别于二进制处理单元, T. Kohonen 所用的处理单元是线性连续的。另外, T. Kohonen 网络还包含了许多邻近的、同时激活的输出节点,即网络的输出不是单个“优胜”的神经元,而是采用相当大数目的一组输出节点来表示输入模式的分类,这样可以便于分析可视图像和语言声谱。

1976年,美国波士顿大学自适应系统研究中心的 S. Grossberg 教授和他的夫人 G. A. Carpenter 共同提出了著名的自适应共振理论(Adaptive Resonance Theory, ART)模型,其学习过程具有自组织和自稳定的特征。S. Grossberg 提出了一个重要的理论,他指出:若在全局神经节点中有一个神经节点特别兴奋,其周围的所有节点将受到抑制。同时, S. Grossberg 还讨论了神经网络的记忆理论,节点的激活值与连接权都会随时间的衰减而衰减,具有“忘却”特性。节点激活值的衰减相当快(短期记忆),而连接权有较长的记忆能力,衰减较慢。随后, S. Grossberg 与 G. A. Carpenter 对自适应共振理论模型做了进一步研究,提出 ARTI、ARTII、ARTIII 三种不同的自适应共振网络模型。三个模型中, ARTI 网络只能处理二值的输入, ARTII 比 ARTI 更为复杂并且能处理连续型输入, ARTIII 网络纳入了生物神经元的生物电-化学反应机制,具有了更强的功能与可拓展能力。

1980年,日本放送协会(Nippon Hōsō Kyōkai, NHK)的福岛邦彦教授研发了有名的

“新认知机”(neocognitron)模型。福岛邦彦在 F. Rosenblatt 提出的感知器模型的基础上,引入了隐含层,从而形成了多层的“新认知机”模型,该模型可实现由高层向低层的可变换的反馈连接以及普通的前馈连接。即使在外部刺激信号停止后,该模型也可以通过抑制性反馈和兴奋性前馈作用,继续实现自组织式的学习过程,从而实现联想式学习和模式识别功能。这种网络模型采用有教师指导的训练方式,它可以在噪声干扰的条件下正确识别包括模式变形情况下的手写 0~9 十个数字。

上述的这些研究成果在现在看来,都是开创性的,具有重要的意义。也正是因为拥有了这些具有巨大科学价值的研究成果,才为神经网络的复兴和发展奠定了坚实的基础。1982 年,以美国物理学家 J. J. Hopfield 发表的著名文章《神经网络和物理系统》(Neural Network and Physical System)为标志,神经网络长达十三年的萧条时期宣告结束。

1.2.3 神经网络研究的复兴与再发展

进入 20 世纪 80 年代后,一度发展兴盛的人工智能及冯·诺依曼型电子计算机科学遇到了瓶颈,它们自身存在的问题及局限性逐渐暴露。逻辑符号对于形象思维、联想记忆、视觉等信息的处理束手无策,这迫使人们不得不重新考虑人工智能与计算机的定位。是什么原因导致能够解决复杂计算问题的电子计算机,却不能完成几岁孩子就能完成的普通的具有经验性的工作。在这样的背景下,人们重新认识到神经网络研究的重要性,神经网络的研究逐渐呈现复苏状态。

1982 年,美国加州理工学院物理学家 J. J. Hopfield 教授做了一项具有开拓性的工作。他在美国科学院院刊上发表了一篇论文,将前人的研究成果及经验进行总结概括,将神经网络的理论分析与动力学系统稳定性分析方法相结合,引入了“计算能量函数”的概念。他将李雅普诺夫(Lyapunov)能量函数引入网络训练这一动态系统中,建立了神经网络稳定性判据,并提出信息存储在网络中神经元的连接上的观点。J. J. Hopfield 教授开拓了神经网络用于联想记忆和组合优化等计算的新途径。这项研究成果具有很大的突破性,为神经网络研究的复苏起到了重要的作用。鉴于他开创性的研究成果,他所提出的网络模型被称为 Hopfield 网络模型。由于该模型与电子电路之间存在着明显的对应关系,因此它非常容易被理解且可以用硬件电路来实现。1984 年, J. J. Hopfield 教授在 Hopfield 网络模型的基础上设计并研制了 Hopfield 网络模型的硬件实现电路,并指出网络模型中的神经元均可用普通的运算放大器来实现,而神经元之间的连接可用电子线路模拟,该方案的提出为神经网络的工程实现指明了道路。他同时进行了神经网络应用研究,成功解决了旅行商(Travelling Salesman Problem, TSP)计算难题,引起了学术界的广泛关注。Hopfield 网络模型取得的这些成果吸引了大批的研究者参与到神经网络的研究中,从而使其研究进入了一个新的兴盛时期。

1983 年,美国的研究员 S. Kirkpatrick、C. D. Gelatt 和 M. P. Vecchi 提出了一种模拟高温物体退火过程来找寻全局最优解的方法,称为模拟退火算法(Simulate Anneal

Arithmetic, SAA), 这种方法可用于组合优化问题的求解。模拟退火算法的研究为之后著名的玻尔兹曼(Boltzmann)机的提出奠定了理论基础。

1984年, 多伦多大学的 G. E. Hinton 教授和霍布金斯大学的年轻学者 T. J. Sejnowsk 合作, 借助统计物理学的概念和方法提出了大规模并行网络学习机, 称为玻尔兹曼机, 并明确提出隐单元的概念。玻尔兹曼机模型的学习规则使用模拟退火算法的思想, 有效地解决了 Hopfield 网络无法克服的局部极小值问题。

1985年, 英国 M. J. D. Powell 教授提出了径向基函数。简单来说, 径向基函数是一个取值仅和到原点距离有关的实值函数, 它可以用于解决多变量差值问题, 该函数的提出为多层前向神经网络的学习提供了一种新的方法。1988年, 英国 D. Broomhead 和 D. Lowe 教授首先将径向基函数应用于神经网络设计, 构成了径向基函数(Radial Basis Function, RBF)神经网络。径向基函数神经网络不仅具有良好的推广能力, 而且避免了像 BP 算法那样烦琐的计算, 使学习能得以快速地实现。并且凭借着其学习收敛速度较快的优势, 可有效避免局部极小值的问题。

1986年, 受到 J. J. Hopfield 教授研究内容的影响, 美国认知心理学家的 D. E. Rumelhart 和 J. L. McClelland 带领其研究小组成员撰写了一本名为《Parallel Distributed Processing: Exploration in the Microstructures of Cognition—PDP》的著作, 即《并行分布式处理》。书中提到了神经网络的三个主要特征: 结构、神经元的激活函数(也称作激励函数或传递函数)以及网络的学习方法, 并针对这三方面分别介绍了不同的网络类型。在该书中, 详尽地分析了具有非线性连续激活函数的多层前馈网络的误差反向传播算法(即 BP 算法), 该算法解决了长期以来缺少有效的权值调整算法的难题, 是当今影响最大的一种网络学习方法。另外, 书中提出的并行分布式处理思想也推动了神经网络的研究进入新的高潮阶段。

1.2.4 神经网络研究的热潮

1987年, 国际电气工程师与电子工程师学会(Institute of Electrical and Electronics Engineers, IEEE)在美国圣地亚哥召开了首届国际神经网络学术会议, 会上宣告了国际神经网络协会(International Neural Network Society, INNS)正式成立, 这标志着神经网络研究在世界范围内掀起了新的热潮。

自1988年起, 正式由国际神经网络学会和国际电气工程师与电子工程师学会联合举办每年一次的国际神经网络学术会议。同年, 由世界上著名的三位神经网络专家——美国波士顿大学的 S. Grossberg 教授、芬兰赫尔辛基技术大学的 T. Kohonen 教授及日本东京大学的甘利俊一教授主持创办的世界第一份神经网络杂志《Neural Networks》问世。随后, IEEE 也成立了神经网络协会, 并于1990年3月开始出版神经网络会刊。之后, 各种学术期刊的神经网络特刊层出不穷。