

张立明 编著

Models and Applications
of Artificial Neural
Networks

复旦大学出版社

人工神经网络的 模型及其应用



73.821
586

人工神经网络的模型及其应用

张立明 编著

复旦大学出版社

(沪)新登字 202 号

责任编辑 林溪波
责任校对 马金宝

人工神经网络的模型及其应用

张立明 编著

复旦大学出版社出版

(上海国权路 579 号)

新华书店上海发行所发行 复旦大学印刷厂印刷

开本 787×1092 1/16 印张 15.5 字数 383,000

1993 年 7 月第 1 版 1993 年 7 月第 1 次印刷

印数 1—2000

ISBN7-309-01041-8/O·121

定价: 7.50 元

内 容 提 要

人工神经网络是目前国际上迅速发展的前沿交叉学科。它是模拟生物神经结构的新型计算机系统,具有广泛的应用前景。

本书从生物背景出发,论述了前馈、反馈、竞争三种基本的人工神经网络模型的结构、理论、稳定性、收敛性及其算法和应用。并介绍了较新的 CNN 网络、随机网络、CMAC、模糊神经网络以及分维等最新成果;最后给出了 VLSI 芯片实现的电路方案以及芯片与计算机组合应用等,在每一章都有一些应用实例。

本书可作为计算机、电子学、信息科学、通信、控制等专业的高年级大学生及研究生的教材,也可供有关专业科技人员参考。

D036/06

序

人脑的工作过程始终是人们渴望了解的秘密, 长期来神经生物学家以自下而上的方法从单个神经细胞到神经网络的刺激响应进行研究, 而神经生理学家则以自上而下的途径研究脑的认知和行为功能, 经过努力逐渐对脑的部分功能有了一定的认识, 并不断提出了一系列的设想和见解, 它们激发了工程设计人员期望获得至今的计算机仍然很难解决而人和动物处理得很好的一类问题的解决。科学家开始建立各种工程的和数学的神经元和神经网络的模型, 并且深入地进行了研究。1940 年 McCulloch 和 Pitts 首先提出了简单阈值逻辑单元的神经元模型, 1960 年前后 Rosenblatt 的感知器(Perceptron), Widrow 和 Hoff 的 Adaline 和 Steinbuch 的学习矩阵等都相继提出并引起了工程学术界的很大兴趣, 特别是感知器模型由于它简单而易于实现更引起了人们的注意, 然而由于 Minsky 和 Papert 证明了感知器不能实现复杂逻辑的判断功能使得神经网络的研究一度趋于低落。

1982 年 Hopfield 关于神经网络集成计算能力文章的发表, 以及早些时期 Grossberg 及 Carpenter 在神经网络自适应共振, Kohonen 在神经网络自组织映照和 Rumelhart 在多层感知器网络及反馈传输算法方面的研究成果促使人工神经网络的研究再次兴起了热潮。

当前人工神经网络的研究主要环绕网络的基本特性和结构, 网络的工程应用以及它的硬件实现等方面进行。通过几年来学术界和工程界的广泛深入的研究探索, 人工神经网络的非线性映照, 学习分类和实时优化等基本特性已成为一种重要的信息处理方法普遍用于各种工程和学科的研究领域。特别是人工神经网络的硬件实现, 已经集中力量进行研究开发, 一些专用的、通用的人工神经网络集成芯片和协处理器已有商品出售, 人工神经网络在工程中的实际运用正在形成。了解和掌握人工神经网络的基本性能和理论正在成为广大工程技术人员的迫切需要。本书作者在国内较早地开展了人工神经网络方面的研究工作, 并开设了硕士研究生的课程。本书是在此基础上编写的, 内容精简扼要, 着重于基本概念和基本理论的叙述。它的出版必将为人工神经网络的理论和技术在我国的普及推广发挥积极的作用。

凌燮亭

1992.9.7

前言

人工神经网络是一门发展十分迅速的交叉学科。虽然对神经网络、神经单元的研究很早就已开始,也取得了不少成果,如描述神经元的著名的 H.H 方程,Hebb 的学习规则等,但这些成果还只限于在生物或生物控制领域中应用。人工神经网络是在生物学中神经网络理论的基础上,简化了复杂结构的偏微分 H.H 方程模型,构成了多个神经元互联结构的网络,研究这个复杂的多维非线性大系统中的信息的分布、传递以及处理的理论和方法,并用于解决实际工程问题。由于这些神经单元在形式上比较一致,十分便于 VLSI 的实现。这门学科的研究使诸如生物学、认知科学、非线性科学等基础学科与计算机、电子学、人工智能、微电子、信息处理、模式识别等工程学科有机地结合起来,并且更具有广泛的应用前景。

国际上,对这一学科的研究已有数十年的历史,自 1987 年后开始形成热点,并成立了相应的神经网络国际学术组织。在我国,于 1990 年也举办了神经网络的学术年会,研究工作在高等学校和研究单位蓬勃开展。近年来人工神经网络不仅在理论上、模型上有所发展,而且在应用上也有较快的发展,特别是这一二年芯片的发展十分迅速,已有不少商品化的芯片问世,这使科技人员、大学生和研究生在自己的研究领域中有可能会接触到人工神经网络、有可能应用人工神经网络这一最新高技术成果来解决自己的问题。为此,不少高等学校已为本科高年级学生和研究生开设了神经网络的课程以及及时介绍这一领域的发展。本书作者也是出于这一考虑,为及时满足这一需要,介绍有关这一领域的最新成果,以当前几种主要模型为线索贯穿全书,期望读者对人工神经网络的发展有比较全面的了解,并能应用这些模型解决实际问题。

本书是编者在三年来为研究生讲授《人工神经网络》课程的讲稿基础上,结合近几年的科学研究成果,并参考国内外有关文献资料编著而成的。书中没有采用过多的数学工具,从几个神经网络的基本模型出发,分析其单元的生物背景、结构、比较强调设计考虑、算法公式及应用方面。

全书分为六章。第一章介绍人工神经网络的概念以及人工神经网络发展的历史,按发展的时间叙述,并介绍一些主要人物的贡献。第二章讨论前馈式的网络,分别介绍 M-P 模型、感知器模型以及反向传播模型,并从线性阈值模型出发,讨论隐单元数的选取,分析其收敛性及算法的改进。第三章讨论反馈式网络,包括全互联离散型的反馈网络,全互联的连续型反馈网络,及局部互联的 CNN 网络,讨论其稳定性,权的设计及应用。第四章讨论自组织网络,比较详细地给出了各个模型的算法、公式及应用。第五章介绍其他形式的模型,一部分模型与生物结构较为相似,如随机网络、遗传算法、分维网络等,另一部分是具有比较实用的网络如 CMAC、模糊神经网络等。第六章简单介绍了 VLSI 硬件的人工神经网络实现方法,及一些通用芯片。

需要说明的是,本书的第二章、第三章、第四章、第六章等章节的部分内容,反映了作者及其研究生和课题组的一些研究成果;第四章的第三节 Fukushima 模型是由瞿东晖编写的;书后的不少参考文献中的公式和方法,由研究生在学习本课程时得到了验证。

在编写本书的过程中，自始至终得到凌燮亭教授的热情关心和指导，他审阅了全部书稿，并提了不少有益的建议；何永保教授对本书的编写给予大力支持，复旦大学神经网络课题组的研究同仁也都给了热情帮助；复旦大学出版社责任编辑林溪波为本书得以出版作了大量的工作，在此向他们表示衷心的感谢！

限于作者水平，且成书仓促，书中错误和欠妥之处在所难免，恳请读者指正。

作 者

1992年6月

目 录

序.....	1
前言.....	1
第一章 概论.....	1
第一节 神经细胞及神经细胞组成的网络.....	1
第二节 人工神经网络的特点.....	3
第三节 人工神经网络的发展史.....	6
一、初始发展期.....	6
二、低潮时期.....	7
三、复兴时期.....	8
四、发展高潮期.....	9
第四节 人工神经网络的类型.....	10
一、神经元变换函数的类型.....	10
二、人工神经网络的结构.....	10
三、学习算法上的分类.....	12
第二章 前馈型人工神经网络.....	13
第一节 线性阈值单元组成的前馈网络.....	13
一、M-P 模型	13
二、感知器组合的神经网络.....	14
三、多层的感知器网络.....	19
第二节 自适应线性单元组成的网络.....	24
一、用于非线性分割的自适应网络.....	25
二、多层 Adaline 网络的学习算法.....	26
三、Adaline 和 Madaline 的应用	28
第三节 非线性变换单元组成的前馈网络.....	32
一、网络的结构与数学描述.....	32
二、B-P 的学习算法.....	34
三、B-P 网络的误差曲面讨论	37
四、算法的改进.....	41
五、B-P 网络的设计考虑.....	43
六、B-P 网络的应用举例.....	47
第三章 反馈式人工神经网络.....	52
第一节 离散的单层反馈网络模型.....	55
一、基本公式.....	55

二、稳定点	56
三、网络的稳定性	57
四、外积型 DHNN 权的设计和讨论	60
五、用其他方法进行权的设计	67
六、记忆容量的讨论	70
第二节 连续的单层反馈网络	75
一、连续的 Hopfield 网络的生物背景和数学模型	75
二、CHNN 方程的解及稳定性讨论	78
三、李雅普诺夫稳定性定理	84
四、Hopfield 的能量函数和稳定性判别	89
五、能量函数与优化计算	93
六、应用举例	97
第三节 细胞神经网络	105
一、CNN 网络的模型	105
二、CNN 网络系统的分析	108
三、非对称模板条件下的系统稳定性	112
四、网络权的设计	115
五、CNN 网络的应用举例	118
第四章 自组织竞争人工神经网络	127
第一节 概述	127
一、竞争	128
二、自组织网络的学习规律	129
第二节 自适应共振理论模型	131
一、ART 的模型结构	131
二、ART 的工作流程	133
三、权的设计及学习	134
四、ART 的实现	133
第三节 自组织映照模型	137
一、自组织映照模型的结构和工作过程	137
二、Kohonen 网络的工作原理	141
三、Kohonen 网络的应用举例	146
第四节 Fukushima 网络模型	146
一、人工神经认知机模型	148
二、具有选择注意力的 Fukushima 神经网络	157
第五章 其他类型的人工神经网络模型	161
第一节 随机神经网络	164
一、Boltzmann 分布和能量函数	164
二、Boltzmann 机模拟学习样本的状态概率	165
三、平均场退火法的随机网络	168

第二节	脑模型联接控制器	170
一、	CMAC 模型的结构	171
二、	网络的工作原理分析	173
三、	CMAC 的学习算法及收敛速度	179
四、	CMAC 在机器人控制中的应用	183
第三节	模糊与人工神经网络	184
一、	模糊的基本概念	185
二、	人工神经网络和模糊系统的比较	193
三、	模糊理论基础上的多级神经网络	199
第四节	分形人工神经网络	206
一、	分形概述	206
二、	结构上的分形人工神经网络	207
三、	信息分形的人工神经网络	209
第五节	遗传算法在人工神经网络中的应用	212
一、	概述	212
二、	遗传算法的流程	213
三、	遗传算法的举例	214
第六章	人工神经网络的实现	217
第一节	概述	217
第二节	全模拟电路实现的神经网络	218
一、	神经元的实现	218
二、	权的设计	218
第三节	全数字电路的人工神经网络	223
第四节	脉冲串作为信号源的混合神经网络	225
第五节	具有虚处理单元的神经网络	230
参考文献		234

第一章 概 论

神经网络是最近发展起来的十分热门的交叉学科,它涉及生物、电子、计算机、数学和物理等学科,有着非常广泛的应用背景,这门学科的发展对目前和未来的科学技术的发展将有重要的影响。

长期以来,人们想方设法了解人脑的功能,用物理可实现系统去模仿人脑,完成类似于人脑的工作。计算机就是采用电子元件的组合来完成人脑的某些记忆、计算、判断功能的物理系统。用机器代替人脑的部分劳动是当今科学技术发展的重要标志。现代计算机中每个电子元件的计算速度为纳秒(ns)级,人脑中每个神经细胞的反应时间只有毫秒(ms)级。这样,似乎计算机的运算能力应为人脑的几百万倍。可是,迄今为止,计算机在解决信息初级加工时如视觉、听觉、嗅觉这类简单的感觉识别上却十分迟钝。人在识别文字、图像、声音等方面的能力大大超过计算机,现代计算机要花几十分钟(min),甚至几小时(h)才能完成的识别任务,人只要零点几秒(s)就可以完成了。同样,在机器人控制、人工智能、思维、直觉等方面,就更无法与人相比了。人们希望去追求一种新型的计算机系统,它既有超越于人的计算能力,又有类似于人的识别、智能、联想的能力。

从人脑的结构看,它是由大量神经细胞组合而成的,这些细胞相互联接,每个细胞完成其某一种基本功能,如兴奋与抑制。从整体看,它们相互整合完成一种复杂的计算机思维活动,这些工作是并行的,有机的关联在一起,这种集体的功能就像用透镜得到图像的傅里叶变换一样,十分迅速,在人的日常生活中,每天都有成千上万的信息需要脑来处理,一个简单的动作如端一杯水,打一个电话,就牵涉到记忆、学习和相位变换等功能,而人却可以不加思索地完成,这就说明需要有一种新型的、类似于人脑结构的系统,来完成那些计算机做起来很慢或很困难的工作。

人工神经网络就是采用物理可实现的系统来模仿人脑神经细胞的结构和功能的系统。它是由很多处理单元有机地联接起来,进行并行的工作,它的处理单元十分简单,其工作则是“集体”进行的,它的信息传播、存贮方式与神经网络相似,它没有运算器、存贮器、控制器这些现代计算机的基本单元,而是相同的简单处理器的组合。它的信息是存贮在处理单元之间的连接上,因而,它是与现代计算机完全不同的系统。

本章主要介绍人工神经网络与生物体实际神经网络的比较、人工神经网络的特点、发展的历史及主要类型。

第一节 神经细胞及神经细胞组成的网络

一个神经细胞的构造如图 1-1 所示,主要包括细胞体、树突、轴突和细胞之间相互关联的突触。

细胞体是由细胞核、细胞浆、细胞膜等组成。在高等动物的神经细胞中,除了特殊的无“轴突”神经元外,一般每个神经元都由胞体的轴丘处发出一根粗细均匀、表面光滑的突起,

• 1 •

9410016

长度从几个 μm 到 1m 左右,称为轴突,它的功能是传出从细胞体来的神经信息。树突为细胞

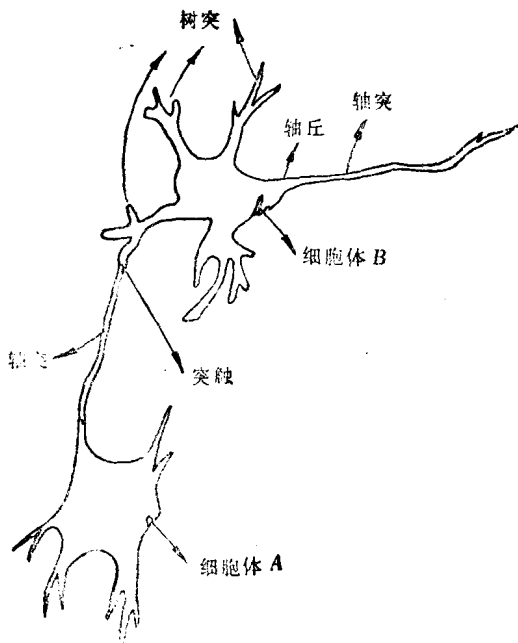


图 1-1 神经细胞的示意图

胞体向外伸出的很多其他突起,它们像树枝一样向四处分散开来,在胞体附近比轴突粗得多,但离开细胞体不远马上就很快分支变细,形成无数粗细不等的树突。它们的作用是向四方收集由其他神经细胞来的信息。信息流是从树突出发,经过细胞体,然后由轴突输出。突触是两个细胞之间连接的基本单元,主要有:

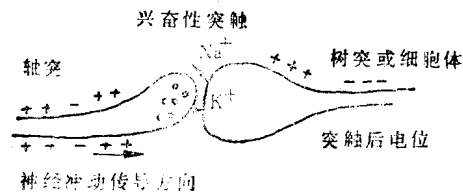


图 1-2 突触联接的示意图

- (1) 一个神经细胞的轴突与另一个神经细胞的树突发生接触;
- (2) 一个神经细胞的轴突与另一个神经细胞的胞体接触。

突触包含两个部分:

一个为突触前成分,表示在轴突的末梢;

一个是突触后成分,为树突的始端,或细胞体与轴突末端接触的部分。突触的联接如图 1-2 所示。

根据突触对下一个神经细胞的功能活动的影响,突触又可分为兴奋性的和抑制性的两种。兴奋性的突触:当神经细胞有冲动传递时,它能使突触后成分去极化,产生兴奋性突触后电位,可能引起下一个神经细胞兴奋。抑制性的突触能使突触后成分超极化,产生抑制突触后电位,使下一个神经细胞抑制。

神经细胞的种类很多,根据功能特性可以分为传入(感觉)神经细胞,中间(联络)神经细胞和传出(运动)神经细胞。

神经细胞单元的信息是宽度和幅度都相同的脉冲串。脉冲串的间隔是随机变化的,如某个神经细胞兴奋,其轴突输出的脉冲串在单位时间内的平均频率高(即发放率高),如神经细胞抑制时,脉冲发放率减少,甚至无脉冲发放。每个神经细胞的输出脉冲发放率是与别的和它相联的神经细胞的发放情况及它们和该神经细胞突触联接情况有关。信息传递在突触处主要是发生化学和电的变化,(这里主要考虑化学变化)即在化学突触的传递过程中,突触前成分囊泡里的神经递质被释放到突触后膜,化学传递的突触传递神经冲动是通过释放神经递质来实现的,当脉冲到来时,贮存在突触囊泡内的神经递质进行排放,这样改变了突触后膜对钠离子(Na^+)、钾离子(K^+)和氯离子(Cl^-)的通透性,使突触后神经细胞相应发生电位变

化。突触传递信息需要一定的延迟时间,对温血动物,延迟时间为0.3~1.0ms。

兴奋性突触在脉冲刺激下,对下一个神经细胞产生兴奋性突触后的电位变化,抑制性突触在脉冲刺激下产生抑制性突触后的电位变化。很多神经细胞通过各自的突触对某一个神经细胞的作用,都形成该神经细胞的后电位变化,电位的变化是可以累加的,该神经细胞后电位是它所有的突触产生的电位总和,当该神经细胞的后电位升高到超过一个阈值,就会产生一个脉冲,从而总和的膜电位直接影响该神经细胞兴奋发放的脉冲数。一般说,每个神经细胞的轴突大约联接100~1000个其他神经细胞,神经细胞的信息就是这样从一个神经细胞传递到另一个神经细胞。

在人脑中大约有140亿个神经细胞单元,根据Stubbz的估计,这些神经细胞被安排进约1000个主要模块内,每个模块有上百个神经网络,每个网络约有10万个神经细胞,信息的传递是从一个神经细胞传到另一个神经细胞,从一种类型的神经细胞传到另一类神经细胞,从一个网络传到另一个网络,有时也从一个模块传到另一个模块。

对于化学突触的信息传递有以下几个原则:

- (1) 只允许脉冲从突触前传向突触后,不允许逆向传递;
- (2) 突触有延迟;
- (3) 突触传递容易产生疲劳;
- (4) 突触前传来的一次冲动常常不足以引起突触后神经细胞产生兴奋,同一个突触传来一系列脉冲,它们刺激间隔比较小,会引起突触后电位的变化,这就是时间上的总和;而很多突触前同时传来的脉冲也能引起突触后电位的变化,这是空间的总和;时间和空间的总和对突触后膜都会产生作用;

- (5) 存在不应期,在不应期内,神经元对刺激不响应。

第二节 人工神经网络的特点

人工神经网络是采用物理可实现的器件或采用现有的计算机来模拟生物体中神经网络的某些结构与功能,并反过来用于工程或其他领域。人工神经网络的着眼点不是用物理器件去完整地复制生物体中神经细胞网络,而是采纳其可利用的部分来克服目前计算机或其他系统不能解决的问题,如学习、识别、控制、专家系统等。随着生物和认知科学的发展,人们对人脑的结构及认知过程了解得越深入,这对人工神经网络的促进作用将会越大,越来越多的生物特性将被利用到工程中去。

图1-3是一个人工神经单元的示意图。图中 $x_1, \dots, x_n$ 表示其他神经元的轴突输出, $w_1, \dots, w_n$ 为其他几个神经元与第 i 个神经元的突触联接, $w_1, \dots, w_n$ 可以是正,也可以是负,分别表示为兴奋性突触和抑制性突触,数值的大小是根据突触不同的化学变化而各不相同。每个人工神经元满足:

$$S_i = \sum_{j=1}^n w_j x_j - \theta_i \quad (1-1)$$

$$u_i = g(s_i) \quad (1-2)$$

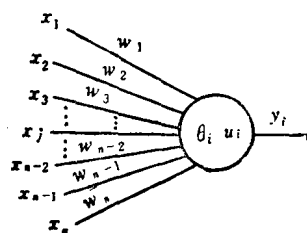


图1-3 人工神经元的示意图

$$y_i = f(u_i) \quad (1-3)$$

其中式(1-1)表示神经元*i*突触后电位的累加值, θ_i 为阈值。式(1-3)中, u_i 为细胞*i*的状态, y_i 为神经单元*i*的输出, $f(u_i)$ 是一个单调上升的函数,当 u_i 大时, y_i 也大,但 $f(u_i)$ 又是一个有限值函数,这是由于在生物体中神经细胞单元脉冲发放率有一个最大值,而不能永远无限上升的缘故。图1-3可以简单地表明上节所述神经细胞的基本功能,但是有以下几点不同:

(1) 在生物体内神经细胞中,来自其他几个神经细胞的轴突输出是发放脉冲串,用脉冲串在单位时间内的发放率来表示其兴奋活动情况,输出的脉冲串引起了突触化学成分的变化,然后引起第*i*个神经细胞突触后电位的变化,后电位的变化又引起 y_i 的脉冲发放率的变化,在人工神经网络中 $x_1, x_2, \dots, x_n$ 和 y_i 大都用电压值来代替, y_i 与 u_i 的关系也是电压与电压的关系。在人工神经网络的硬件实现时,也有人用数字脉冲频率的大小表示 y_i 与 u_i ,但与真正的神经网络还有较大的差距。

(2) 由于(1-1)式、(1-2)式和(1-3)式描写的方程,其信息的传递是采用电压的方式,因而大多数人工神经网络只有空间累加而没有时间累加。

(3) 在人工神经网络中,都没有考虑突触原则中的疲劳和不应期。

(4) 根据方程式(1-1),由于 x_j 为电压,因而权 w_j 为导纳量纲, θ_i 和 s_i 为电流量纲, u_i 和 y_i 为电压量纲,这样,突触在人工神经网络中只要用一个电阻或相当于电阻的电子器件就可以实现。这种简单的权的表示与真正的生物突触中的化学反应和电反应也有很大的差距。

(5) 神经细胞的种类很多,但在一类人工神经网络中,神经元的种类通常仅为一种,少数有2~3种。

(6) 神经网络是由大量的神经细胞组成,而且神经细胞会不断死亡和增殖,在人工神经网络中,由于物理器件的限制,它的神经单元的数目大大小于真正神经网络中的细胞数,因而其功能也比真正神经网络差得多。这是两者最大的区别。当然,随着新器件的发展(如分子器件)会逐步缩小这个差距。

虽然人工神经网络与真正的神经网络有上述这些差别,但它与目前冯·诺伊曼计算机相比,由于它吸取了生物神经网络的部分优点,因而有其固有的特点:

(1) 人工神经网络在结构上与目前的计算机根本不同,它是由很多小的处理单元互相联接而成,每个处理单元的结构如图1-3所示,并由方程(1-1)、(1-2)、(1-3)来描述,每个单元的功能简单,但大量简单的处理单元集体的、并行的活动得到预期的识别、计算的结果,具有较快的速度。

(2) 人工神经网络具有很强的容错性,即局部的或部分的神经元损坏后,不影响全局的活动。

(3) 人工神经网络所记忆的信息是存贮在神经元之间的权中,从单个权中看不出其贮存信息的内容,因而是分布式的存贮方式。

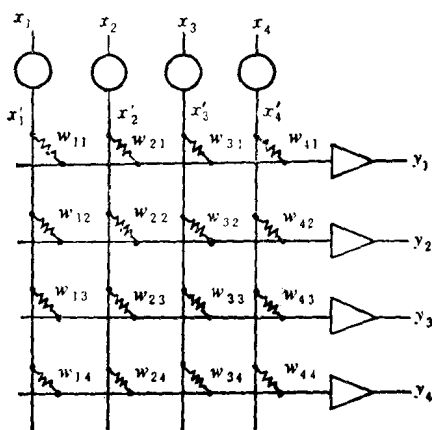


图1-4 简单的人工神经网络

(4) 人工神经网络具有十分强的学习功能, 人工神经网络中的连接权和连接的结构都可以通过学习而得到。

现在, 我们以图 1-4 所示的一个简单的人工神经网络的例子来说明上述的几个特点。图中, x_1, x_2, x_3, x_4 为输入, 经感知细胞输出为 x'_1, x'_2, x'_3, x'_4 , 然后经过突触权 w_{ij} 加到 y_1, y_2, y_3, y_4 的输入端, 进行累加。为了简单起见, 我们把式(1-1), 式(1-2)和式(1-3)的表达式改写为:

设: $\theta_i=0$, 则有

$$s_i = \sum_{j=1}^n w_{ij} x'_j \quad (1-1)'$$

$$u_i = s_i \cdot 1 \quad (\text{量纲变换}) \quad (1-2)'$$

$$y_i = f(u_i) = \begin{cases} 1 & u_i \geq 0 \\ -1 & u_i < 0 \end{cases} \quad (1-3)'$$

又设输入 x'_j ($j=1, \dots, 4$) 也为二值变量(即 $x'_j = \pm 1$), 且: $x'_j = x_j$,

x_j 是感知器的输入, 我们用矢量 $x^1 = [1, -1, -1, 1]^T$ 表示眼看到花、鼻嗅到花香的感知输入, 从 $x^1 \rightarrow y^1$ 可通过一个联接矩阵 W_1 来得到。

$$W_1 = \begin{bmatrix} -0.25 & +0.25 & +0.25 & -0.25 \\ -0.25 & +0.25 & +0.25 & -0.25 \\ +0.25 & -0.25 & -0.25 & +0.25 \\ +0.25 & -0.25 & -0.25 & +0.25 \end{bmatrix} \quad (1-4)$$

$$y^1 = f(W_1 x^1)$$

经计算:

$$y^1 = [-1, -1, +1, +1]^T$$

这表示网络决策 x^1 为一朵花的概念。

我们看到从 $x^1 \rightarrow y^1$ 的变化不是经过串型的计算得到的, 因为矩阵 W_1 是一个可以用 VLSI 中的电阻矩阵来实现的, 而 $y_i = f(v_i)$ 也可用一个简单的运算放大器来模拟, 不管 x^1 和 y^1 的维数如何增加, 整个计算都只用了一个运放的转换时间, 网络的动作是并行的。

如果 $x^2 = [-1, +1, -1, +1]^T$, 表示眼看到炸猪排、鼻嗅到炸猪排的香味的感知器输入, 通过矩阵

$$W_2 = \begin{bmatrix} +0.25 & -0.25 & +0.25 & -0.25 \\ -0.25 & +0.25 & -0.25 & +0.25 \\ -0.25 & +0.25 & -0.25 & +0.25 \\ +0.25 & -0.25 & +0.25 & -0.25 \end{bmatrix} \quad (1-5)$$

得到 $y^2 = [-1, +1, +1, -1]^T$, 表示网络决策 x^2 为炸猪排。

从式(1-4), (1-5)的权来看, 我们并不知道其输出结果是什么, 从局部权的分布看, 也很难看出 W 中存贮什么, 这是因为信息是分布式的存贮在权中, 如果把(1-4)、(1-5)式相加, 得到一组新的权:

$$W = W_1 + W_2 = \begin{bmatrix} 0 & 0 & 0.5 & -0.5 \\ -0.5 & 0.5 & 0 & 0 \\ 0 & 0 & -0.5 & 0.5 \\ 0.5 & -0.5 & 0 & 0 \end{bmatrix} \quad (1-6)$$

由 x^1 输入到这网络时,通过矩阵运算得到 y^1 , 由 x^2 输入得到 y^2 , 这说明 W 中存贮了两种信息,当然还可存贮多种信息,只要用多个已知的例子进行学习,在权中就可以存贮多种信息。

如果输入感知器中损坏了一个,设第三个损坏了,则 $x^1=[1,-1,0,1]^T$, 经 W 算得 $y^1=[-1,-1,+1,+1]^T$, 而 $x^2=[-1,+1,0,1]^T$ 经 W 算得 $y^2=[-1,+1,+1,-1]^T$, 结果和前面的一样。这说明人工神经网络有一定的容错能力。

从上面的例子可以看出人工神经网络不像现在的计算机,有存贮单元、运算单元和控制单元,它的存贮单元和运算控制单元都融合在一个网络中,它的结构,工作步骤及方式完全与现代计算机不同,表 1-1 简单地列出了它们的差别,这里还要指出,人工神经网络与一般并行计算机或并行处理器的算法也不相同,因为并行机是多个 CPU 处理器并行工作,它的处理单元比人工神经元要复杂得多,它的算法和软件更比人工神经网络复杂得多。并行机只能解决计算机的处理速度问题,并没有上述的人工神经网络的特点和集体动作的功能。

表 1-1

	现 代 计 算 机	人 工 神 经 网 络 (ANN)
元件间的连接	2~3	100~1000
实现结构	前馈式加很少反馈	前馈式 全反馈式
功 能	不能学习 串行、编程	能学习 并行、不编程或简单编程
记忆方式	存贮器 集中记忆	分布式记忆
容 错	元件损坏 无法工作	元件损坏能继续工作 或学习后继续工作
应 用	计算 逻辑判断 信息处理	识别感知 智能控制 专家系统等

第三节 人工神经网络的发展史

人工神经网络的发展可追溯到一个世纪前,根据文献^[17],可把这一历史分为四个时期。第一时期为初始发展期,自 1890 年至 1969 年;第二时期为低潮时期,自 1969 年至 1982 年;第三时期为复兴时期,自 1982 年至 1986 年;第四时期为发展高潮期,自 1986 年到现在。下面就每个时期的情况及有贡献的人物略作介绍。

一、初始发展期

1890 年,美国生理学家 W. James 出版了《生理学》^[2]一书,该书首次阐明了有关人脑结构及其功能,以及一些相关学习、联想记忆的基本规则,并指出:人脑中,当两个基本处理单元同时活动,或两个单元靠得比较近时,一个单元的兴奋会传递到另一个单元,而且一个单元的活动程度与它周围单元的活动数目和活动的密度是成正比的。这些论述有些是该书作者的推断,但在以后的发展中都证明这些推断是正确的。

大约经过半个世纪以后,McCulloch 和 Pitts 发表了十分有名的论文^[3]。在这篇文章

中,他们用已知的神经细胞生物基础,描述了一个简单的人工神经元模型,这个神经元的活动是服从二值变化的,即兴奋和抑制,任何兴奋突触有输入激励后,使神经元兴奋,与过去神经元活动情况和神经元的位置无关,任何抑制突触被输入激励后,这个神经元即被抑制。突触的值是不变的,突触存在延迟时间为 0.5ms。他们描述的模型后来被称为 M-P 模型。现在看来,这个模型是过于简单了,但他们的贡献在于:

- (1) 证明了用 M-P 模型能完成任意有限的逻辑运算;
- (2) 他们是在 W. James 后第一个采用集体并行计算结构来描述人工神经元和网络工作的;
- (3) 他们的工作为以后进一步的研究提供了依据。

1949 年 Donald O. Hebb 发表论著《行为自组织》^[6],首先定义了一种调整权的方法,称为 Hebbian,他指出当一个 A 细胞的轴突,充分靠近细胞 B,并持续不断地激励它,它们两个细胞的效应都增长,即 B 的活动增加了, A 的活动也增加了。Hebb 提出了很多有价值的观点,这对以后人工神经网络的结构和算法都有很大的影响,具体有以下几点:

- (1) 他提出了在一个神经网络里信息是贮藏在突触联接的权中;
- (2) 联接权的学习律是正比于两个被联神经细胞的活动状态值的乘积;
- (3) 假设权是对称的,从 A 细胞到 B 细胞的权等于 B 细胞到 A 细胞的权;
- (4) 细胞的互相联接的结构是它们权的改变创造出来的,例如一组弱联接的神经细胞同时活动,则细胞间的联系加强,形成强联接。直到现在,Hebb 的学习算法仍在不少人工神经网络中应用。

1958 年 Frank Rosenblatt 定义了一个神经网络结构^[8],称为感知器 (perceptron),这是第一个真正的人工神经网络,因为他在 IBM704 计算机上进行了模拟,从模拟结果看,perceptron 是有能力通过调整权的学习达到正确分类的结果,因此它可以称为一个学习机。Rosenblatt 用 Perceptron 来模拟一个生物视觉模型,输入是由随机的一组细胞组成,代表视网膜的一个小的范围,每个细胞又与下一层的神经细胞(称 AU 单元)相联,而 AU 又与第三层(RU)相联, RU 为输出层, Perceptron 的目的是通过学习,使对应的输入模板有正确的 RU 输出。初始的 Perceptron 学习机制是自组织的,因而响应的发生与初始的随机值有关,后来,他也加入一些教师进行训练,这些与后来在反向传输算法和 Kohonen 的自组织算法类似,因此 Rosenblatt 的思想有相当的活力。

1960 年 Bernard Widrow 和 Marcian Hoff 发表了“自适应开关电路”的论文^[9],他们从工程的观点出发,不仅在计算机上模拟了这种神经网络,而且还做成了硬件,他们介绍的器件是一个称为“Adaline”的累加输出单元,输出的值是 ± 1 的二值变量,权在 Widrow 和 Hoff 的文章中称为增益,他们主要提出了 Widrow-Hoff 算法,使增益的学习速度较快,而且还有较高的精度,后来这个算法被称为 LMS 算法,在数学上就是人们所知的速降法。Widrow-Hoff 的算法不需要微分,其权的变化是正比于实际输出值与要求输出值之间的差和输入信号的符号,这种算法在以后反传网络(Back-Propagation)算法和其他信号处理系统中应用十分广泛。

二、低潮时期

1969 年 Marrin Minsky 和 Seymour Papert 发表了名为“Perceptrons”的论著^[7],该书