

高等学校教材

神经网络

——理论、技术、方法及应用

赵庶旭 党建武 张振海 张华卫 编

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

神经网络—— 理论、技术、方法及应用

赵庶旭 党建武 张振海 张华卫 编

中国铁道出版社

2013年·北京

内容简介

本书主要对目前神经网络领域的理论、主流的技术方法和开发应用进行了系统的归纳和阐述。全书共分9章,分别介绍了绪论、神经网络基本模型、神经网络学习理论、前馈型神经网络、反馈神经网络、模糊神经网络、脉冲耦合神经网络、智能算法和神经网络集成。

本书可作为计算机科学与技术、自动控制、信号与信息处理等专业本科生和研究生教材,也可作为相关工程技术及研发人员的参考书。

图书在版编目(CIP)数据

神经网络:理论、技术、方法及应用/赵庶旭等编. —
北京:中国铁道出版社,2013.5
ISBN 978-7-113-16383-9

I. ①神… II. ①赵… III. ①人工神经网络 IV.
①TP183

中国版本图书馆 CIP 数据核字(2013)第 087298 号

书 名: 神经网络——理论、技术、方法及应用
作 者: 赵庶旭 党建武 张振海 张华卫 编

策 划: 刘红梅

责任编辑: 刘红梅 吕继函

编辑部电话: 010-51873133

电子信箱: mm2005td@126.com

封面设计: 冯龙彬

责任校对: 焦桂荣

责任印制: 李 佳

出版发行: 中国铁道出版社(100054,北京市西城区右安门西街8号)

网 址: <http://www.51eds.com>

印 刷: 航远印刷有限公司

版 次: 2013年5月第1版 2013年5月第1次印刷

开 本: 787mm×1092mm 1/16 印张: 12.75 字数: 320千

书 号: ISBN 978-7-113-16383-9

定 价: 30.00元

版权所有 侵权必究

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社读者服务部联系调换。

电 话: 市电(010)51873170,路电(021)73170(发行部)

打击盗版举报电话: 市电(010)63549504,路电(021)73187

前 言

20 世纪 80 年代后期,神经网络的研究开发在欧美及日本等工业发达国家掀起了热潮。随之而来,我国在该方向上的研究与应用也是多点开花,取得了令人瞩目的诸多成果。对神经网络理论、方法及技术应用的研究,多年来一直是相关学科领域理论研究和多行业内应用研究的持续热点,尽管随后在对神经网络的研究过程中出现过低潮,但随着在学习理论、网络模型、高性能计算等方面许多新概念的提出,神经网络领域内的研究依然存在着持久不衰的新鲜活力。

神经网络作为发展迅速的交叉学科,涉及到生物学、医学、心理学、认知学、信息论、数学、计算机科学和微电子技术等多种学科。编者根据自己多年从事神经网络和智能算法方向上的科学研究,结合在铁路智能控制、图像处理领域的应用工程实践经验和相关的教学经验,有重点地进行了本书的编写工作。

本书在编写过程中,参考了大量书籍,目的在于使读者通过本书能够较为全面、系统地对神经网络的基本原理和理论,对主流的可行性高的技术方法,以及目前研究的焦点进行把握。

全书以神经网络的理论、方法、技术和实际应用为主线进行内容安排,主要内容如下:第 1 章为绪论,简单阐述神经网络的历史发展、应用领域和发展方向;第 2 章介绍了人工神经元,完成对神经网络基础模型的介绍;第 3 章对神经网络中的知识的表示方法进行介绍,并说明目前神经网络中采用的主要学习算法;第 4 章和第 5 章从神经网络的网络结构角度,对前馈型神经网络和反馈神经网络的典型模型进行介绍,并针对不同的模型说明其特征和学习算法,重点是对这些神经网络的数学分析、训练方法、性能优化等方面进行系统地、细致地分析和阐述,并结合应用实例进行讲解;第 6 章是在系统介绍模糊神经网络理论的基础上,详细讨论模糊神经网络系统,介绍针对铁路运输调度系统中的列车行车问题,运用常见的模糊网络模型来设计求解的方法;第 7 章在介绍视觉系统的工作原理及典型数学描述模型的基础上,引出脉冲耦合神经网络模型,并对脉冲耦合神经网络的基本模型、运行机理及基本特性进行系统地介绍;遗传算法、模拟退火算法、进化算法及禁忌搜索称作指导性搜索法,而神经网络混沌搜索则属于系统动态演化方法,第 8 章将这些智能方法的原理和方法及神经网络融合来解决最优化问题;第 9 章从理论和实现方法上对神经网络集成进行介绍。目前,神经网络集成已被视为一种有广阔应用前景的工程化神经计算技术,已经成为机器学习和神经计算领域

的研究热点,神经网络集成模型不仅有助于加深对机器学习和神经计算的深入研究,还有助于提升工程中问题的解决能力。

本书由赵庶旭、党建武、张振海、张华卫编写,其中,第2、6章由赵庶旭编写,第1、9章由党建武编写,第5、7、8章由张振海编写,第3、4章由张华卫编写。

限于编者水平有限,编写过程中错误、遗漏及不妥之处敬请各位同行、专家、读者不吝指正。

编 者

2013年2月

目 录

第 1 章 绪 论	1
1.1 人工神经网络发展	1
1.2 人工神经网络发展及应用	3
1.3 生物学的启示	4
复习思考题	5
第 2 章 神经网络基本模型	6
2.1 神经网络	6
2.2 人工神经元模型及表示方法	8
2.3 网络结构	15
复习思考题	17
第 3 章 神经网络学习理论	18
3.1 神经网络的知识表示	18
3.2 神经网络的学习理论	22
复习思考题	37
第 4 章 前馈型神经网络	38
4.1 线性阈值单元组成的前馈网络	38
4.2 非线性变换单元组成的前馈网络	49
4.3 径向基函数神经网络	58
4.4 应用举例	70
复习思考题	76
第 5 章 反馈神经网络	77
5.1 离散的 Hopfield 神经网络	77
5.2 联想记忆	78
5.3 连续型 Hopfield 神经网络	79
5.4 A/D 转换网络	81
5.5 Hopfield 神经网络用于求解组合优化问题	82

5.6 应用举例	83
复习思考题	87
第6章 模糊神经网络	88
6.1 模糊神经网络理论	88
6.2 应用神经网络构造模糊控制系统	97
6.3 应用案例	105
6.4 模糊神经网络求解列车运行安全模糊控制问题	112
复习思考题	127
第7章 脉冲耦合神经网络	128
7.1 视觉系统及其模型	128
7.2 脉冲耦合神经网络基本模型	130
7.3 脉冲耦合神经网络的理论基础	134
7.4 脉冲耦合神经网络的应用	138
复习思考题	142
第8章 智能算法	143
8.1 禁忌搜索算法	143
8.2 模拟退火算法	156
8.3 遗传算法	165
复习思考题	188
第9章 神经网络集成	189
9.1 神经网络集成的基本原理	189
9.2 集成方法	190
9.3 集成结论的生成	191
9.4 个体的生成	192
9.5 研究发展方向	194
复习思考题	195
参考文献	196

第1章 绪 论

当我们产生阅读、呼吸、运动和思考等任意行为的时候,就开始使用复杂的生物神经网络系统。该神经网络系统由约 10^{11} 个神经元的高度互连的集合组成,帮助我们进行阅读、呼吸、运动和思考。人脑中的每一个生物神经元都是生物组织和化学物质的有机结合。从神经元处理事务的原理来看,类似于复杂的微处理器,而其神经结构则分为与生俱来的和后期训练养成的两种。

目前,科学家们对生物神经网络工作机理的认识属于快速发展阶段。一般认为,包括记忆功能在内的所有生物神经功能,都存储在神经元和神经元之间的连接上;学习功能可以理解为是在经验的基础上,在神经元之间建立新的连接或对已有的连接进行修改的过程。根据以上关于神经机理的认识,我们是否可以思考:既然我们已经对生物神经网络有一个基本的认识,那么能否利用一些简单的人工“神经元”构造一个小系统,然后对其进行训练,从而使它们具有一定有用功能呢?回答是肯定的。本书正是要讨论有关人工神经网络工作机理的一些方法和技术。

这里考虑的神经元不是生物神经元,它们是对生物神经元极其简单的抽象,可以用程序或硅电路实现。虽然由这些神经元组成的网络的能力远远不及人脑的那么强大,但是可对其进行训练,以实现一些有用的功能。本书所要介绍的正是有关于这样的神经元,以及包含这些神经元的网络及其训练方法。

1.1 人工神经网络发展

在人工神经网络的发展历程中,涌现了许多在不同领域中富有创造性的传奇人物,他们艰苦奋斗几十年,提出了许多至今仍然让我们受益的概念。许多作者都记载了这一历史。一本特别有趣的书是由 John Anderson 和 Edward Rosenfeld 撰写的《神经计算:研究的基础》,在该书中,他们收集并编辑了一组由 43 篇具有特别历史意义的论文,每一篇前面都有一段历史观点的导言。

我们简单地回顾一下神经网络的主要发展历史。

对技术进步而言,有两点是必需的:概念与实现。首先,必须有一个思考问题的概念,根据这些概念明确所面临的问题。这就要求概念包含一种简单的思想,或者更具特色,并且引入数学描述。为了理解这一点,让我们看看心脏的研究历史。在不同时期,心脏被看成灵魂的中心或身体的热源。17 世纪的医生们认识到心脏是一个血泵,于是科学家们开始设计实验,研究泵的行为。这些实验最终开创了循环系统理论。可以说,没有泵的概念,就不会有人们对心脏的深入认识。

概念及其相应的数学描述还不足以使新技术走向成熟,除非能通过某种方式实现这种系统。比如,虽然多年前就从数学上知道根据计算机辅助层析成像(CAT)扫描可以重构图像,但

是直到有了高速计算机和有效的算法才使其走向实用,并最终实现了有用的 CAT 系统。

神经网络的发展史同时包含了概念创新和实现开发的进步。但是这些成果的取得并不是一帆风顺的。

神经网络领域研究的背景工作始于 19 世纪末和 20 世纪初。它源于物理学、心理学和神经心理学的跨学科研究,主要代表人物有 Herman Von Helmholtz 等。这些早期研究主要还是着重于有关学习、视觉和条件反射等一般理论,并没有包含有关神经元工作的数学模型。

现代对神经网络的研究可以追溯到 20 世纪 40 年代 Warren McCulloch 和 Walter Pitts 的工作。他们从原理上证明了人工神经网络可以计算任何算术和逻辑函数。通常认为他们的工作是神经网络领域研究工作的开始。

在 McCulloch 和 Pitts 之后,Donald 指出,经典的条件反射是由单个神经元的性质引起的,提出了生物神经元的一种学习机制。

人工神经网络第一个实际应用出现在 20 世纪 50 年代后期, Frank Rosenblatt 提出了感知机网络和联想学习规则。Rosenblatt 和他的同事构造了一个感知机网络,并公开演示了它进行模式识别的能力。这次早期的成功引起了许多人对神经网络研究的兴趣,不幸的是,后来研究表明基本的感知机网络只能解决有限的几类问题。

同时, Bernard Widrow 和 Ted Hoff 引入了一个新的学习算法用于训练自适应线性神经网络。它在结构和功能上类似于 Rosenblatt 的感知机。Widrow - Hoff 学习规则至今仍然还在使用。但是, Rosenblatt 和 Widrow 的网络都有同样的固有局限性。这些局限性在 Marvin Minsky 和 Seymour Papert 的书中有广泛的论述。Rosenblatt 和 Widrow 也十分清楚这些局限性,并提出了一些新的网络来克服这些局限性,但是他们没能成功找到训练更加复杂网络的学习算法。

许多人受到 Marvin Minsky 和 Seymour Papert 的影响,相信神经网络的研究已走入了死胡同。同时由于当时没有功能强大的数字计算机来支持各种实验,从而导致许多研究者纷纷离开这一研究领域。神经网络的研究就这样停滞了十多年。

即使如此,在 20 世纪 70 年代,科学家们仍然在该领域开展了许多重要的工作。1972 年 Teuvo Kohonen 和 James Anderson 分别独立提出了能够完成记忆的新型神经网络。这一时期, Stephen Crossberg 在自组织网络方面的研究也十分活跃。

前面我们说过,在 20 世纪 60 年代,由于缺乏新思想和用于实验的高性能计算机,曾一度动摇了人们对神经网络的研究兴趣。到了 80 年代,随着个人计算机和 workstation 计算能力的急剧增强和广泛应用,以及不断引入新的概念,克服了摆在神经网络研究面前的障碍,人们对神经网络的研究热情空前高涨。

有两个新概念对神经网络的复兴具有极其重大的意义。一个是用统计机解释某些类型的递归网络的操作。这类网络可作为联想存储器。物理学家 John Hopfield 的研究论文论述了这些思想。另一个是在 20 世纪 80 年代,几个不同的研究者分别开发出了用于训练多层感知机的反传算法。其中最具影响力的反传算法是 David Rumelhart 和 James McClelland 提出的,该算法有力地回答了 20 世纪 60 年代 Minsky 和 Papert 对神经网络的责难。

人工神经网络已经成为智能科学的一个重要分支,同时多个领域的新进展对神经网络研究领域重新注入了活力。要解决人类在 21 世纪所面临的许多困难,诸如能源的大量需求、环境的污染、资源的耗竭、人口的膨胀等,单靠现有的科学成就是很不够的,必须向生物学习,寻找新的科技发展的道路。

自20世纪80年代以来,神经网络的研究取得了一定进展,构造了许多神经网络模型,关于神经计算机系统的研究工作在最近也出现了进展。但是,对于神经网络理论、方法的研究依然属于起步阶段。目前,在该领域内出现了一些新的主要发展方向:

- (1)脉冲耦合神经网络。
- (2)神经元集群网络模型。
- (3)非线性动力学模型。
- (4)记忆模式。
- (5)人工脑。
- (6)神经计算机。

神经网络具有特有的信息处理能力,成功的实现了神经网络专家系统、模式识别、智能控制、求解组合优化问题等,显示了神经网络的潜力。不断拓宽神经网络的应用领域,将会促进神经网络的发展。神经网络与传统人工智能、模糊数学、子波分析等的结合,将给智能科学和技术的发展提供新方法、新途径。

人类在探索宇宙空间、基本粒子、生命起源等科学领域的进程中经历着艰辛的道路。从智能科学的高度,通过与认知神经科学、计算神经科学、认知科学等结合,将人脑神经网络与人工神经网络结合,探索智能的本质,神经网络的研究将取得突破性进展,开创神经网络研究的新纪元。

1.2 人工神经网络发展及应用

神经网络适合于解决实际问题,其应用领域在不断扩大,它不仅可以广泛应用于工程、科学和数学领域,也可广泛应用于医学、商业、金融和文学等领域。神经网络在许多领域的广泛应用,使其极具吸引力。同时,可以用神经网络解决过去许多计算量很大的复杂工业问题。神经网络已被用于许多领域,下面列举一些其在不同行业的应用。

(1) 航空

高性能飞行器自动驾驶仪,飞行路径模拟,飞机控制系统,自动驾驶优化器,飞行部件模拟,飞行器部件故障检测器。

(2) 汽车

汽车自动导航系统,驾驶行为分析器,卡车制动器诊断系统,车辆调度,运送系统。

(3) 国防

武器操纵,目标跟踪,目标辨识,面部识别,新型的传感器,声纳、雷达和图像信号处理(包括数据压缩、特征提取、噪声抑制、信号/图像的识别)。

(4) 电子

代码序列预测,集成电路芯片布局,过程控制,芯片故障分析,机器视觉,语音综合,非线性建模。

(5) 娱乐

动画,特技,市场预测。

(6) 金融

不动产评估,借贷咨询,抵押审查,公司证券分级,投资交易程序,公司财务分析,通货价格预测;支票和其他公文阅读器,信贷申请的评估器;政策应用评估,产品优化。

(7) 制造

生产流程控制,产品设计和分析,过程和机器诊断,实时微粒识别,可视质量监督系统,啤酒检测,焊接质量分析,纸张质量预测,计算机芯片质量分析,磨床运转分析,化工产品设计分析,机器性能分析,项目投标,计划和管理,化工流程系统动态建模。

(8) 医疗

乳房癌细胞分析,EEG 和 ECG 分析,修复设计,移植次数优化,医院费用节流,医院质量改进,急诊室检查建议。

(9) 石油和天然气

石油和天然气的探查。

(10) 机器人

轨道控制,铲车机器人,操作手控制器,视觉系统。

(11) 语音

语音识别,语音压缩,语音识别,文本到语音的综合。

(12) 电信

图像和数据压缩,自动信息服务,实时语言翻译,客户支付处理系统。

从上述在不同领域中的应用,可以发现神经网络应用的数量,投入到神经网络软硬件上的资金和公众对这些设计的兴趣都在快速增长。

1.3 生物学的启示

本书所讲的人工神经网络和与它对应的生物神经网络有很大区别。本节我们将简单介绍人脑功能中那些对人工神经网络研究有启示的特征。

人脑由大量(约 10^{11} 个)高度互连的单元(每个单元约有 10^4 个连接)组成,这些单元被称为神经元。就研究的目的来看,这些神经元由 3 部分组成:树突、细胞体和轴突。树突是树状的神经纤维接收网络,它将电信号传送到细胞体,细胞体对这些输入信号进行整合并进行阈值处理。轴突是单根长纤维,它把细胞体的输出信号导向其他神经元。一个神经细胞的轴突和另一个神经细胞树突的结合点称为突触。神经元的排列和突触的强度(由复杂的化学过程决定)确立了神经网络的功能。两个生物神经元的简化图如图 1.1 所示。

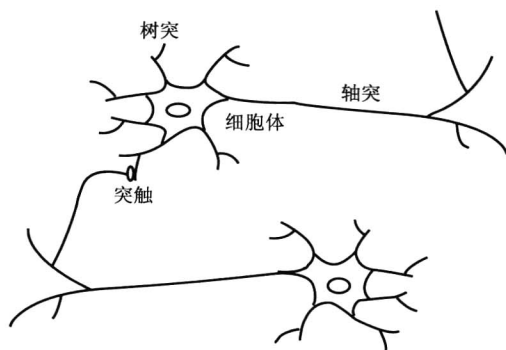


图 1.1 两个生物神经元的简化图

一些神经结构是与生俱来的,而其他部分则是在学习的过程中形成的。在学习的过程中,可能会产生一些新的连接,一些连接也可能会消失,这个过程在生命早期最为显著。比如,如果在某一段关键的时期内禁止一只小猫使用它某一只眼睛,则它的这只眼在以后很难形成正常的视力。

神经结构在整个生命期内不断地进行着改变,后期的改变主要是加强或减弱突触连接。例如,现在已经确认,新记忆的形成是通过改变突触强度而实现的。所以,认识一位新朋友面孔的过程中包含了各种突触的改变过程。

人工神经网络却没有大脑那么复杂,但它们之间有两个关键相似之处。首先,两个网络的构成都是可计算单元的高度互连(虽然人工神经元比生物神经元简单得多)。其次,处理单元之间的连接决定了网络的功能。本书的根本目标就是在人工神经网络中采用合适的连接来解决特定的问题。

值得注意的是,虽然生物神经元相对于电子电路来说非常慢(即 10^{-3}s 比 10^{-9}s 用时更长),人脑却能以比现有计算机快得多的速度完成许多任务。这主要是因为生物神经网络具有巨大的并行性,即所有的神经元能同时操作。即使大多数人工神经网络是在传统的数字计算机上实现的,但并行处理结构使它们适合于采用 VLSI、光学器件和并行处理技术实现。

复习思考题

1. 了解神经网络的发展。
2. 了解人工神经网络的应用领域。

第2章 神经网络基本模型

对生物神经元特别是人脑的认识,促进了对人工神经网络的研究,逐步形成了对人工神经元的数学描述,并据此形成了使用最为广泛的前馈和反馈网络基础模型,最终演化为具有优秀计算能力的各类人工神经网络模型。本章重在通过神经元模型说明人工神经网络节点的计算机理和基础神经网络模型的描述方法。

2.1 神经网络

2.1.1 神经网络定义

自从认识到人脑的计算与传统的数字计算机相比是完全不同的方式开始,关于人工神经网络(以下简称“神经网络”)的研究工作就开始了。人脑是一个高度复杂的、非线性的和复杂的计算机器(信息处理系统)。人脑能够通过它最基础的组成成分——神经元,进行比今天已有的、最快的计算机还要快许多倍速度的特定计算(如模式识别、感知和运动神经控制)。例如人类的视觉系统功能为我们提供一个关于周围环境的描述,同时也提取了人和环境交互所需的信息。

以蝙蝠的声纳定位为例,蝙蝠的声纳是一个活动回声定位系统,可以搜集目标的相对速度、目标大小、目标不同特征的大小及它的方位角和仰角的信息。所有信息都从对目标的回声中提取,而所有需要的复杂神经计算只在“李子”般大小的脑中完成,最终以很高的成功率灵巧地完成追逐和捕捉目标。

那么,人脑或蝙蝠的脑是如何做到这一点的呢?人和动物一出生就有精巧的构造,同时具备了我們通常称为“经验”而建立其自己规则的潜力。确实,经验是经时间积累的,出生后人脑在前两年内发生了最戏剧性的发展(即硬连接),但是发展将超越这个阶段并继续进行。

最普遍形式的神经网络就是通过电子器件或软件编程对人脑完成特定任务或感兴趣功能的建模。在本书中,我们主要介绍重要的神经网络。为了获得好的网络计算结果,神经网络是一个很庞大的简单计算单元间的相互连接(这些简单计算单元称为“神经元”)。据此我们给出将神经网络看作一种自适应机器的定义:一个神经网络是一个由简单处理元构成的规模宏大的并行分布式处理器,天然具有存储经验知识和使之可用的特性。人工神经网络与人脑在两个方面具备相似性,即

(1)知识的获取依赖于外界环境中学习。

(2)通过神经元互联的连接强度,即突触权值,来对知识储存。

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。神经网络在机器学习和智能计算相关的文献中也称为神经计算机、连接主义网络、并行分布式处理器等。本书一律使用“神经网络”这个术语,偶尔也用“神经计算机”或“连接主义网络”。

神经网络的计算能力很明显有以下两点:大规模并行的分布式结构;神经网络学习能力及由此而来的泛化能力。泛化是指神经网络对不在训练(学习)集中的数据可以产生合理的输

出。这两种信息处理能力让神经网络可以解决一些当前还不能处理的复杂的(大型)问题。但是在实践中,神经网络不能单独做出解答,它们需要被整合在一个协调一致的系统工程方法中。具体地讲,一个复杂问题往往被分解成若干个相对简单的任务,而神经网络处理与其能力相符的子任务。

2.1.2 神经网络的性质和能力

神经网络本身结构和计算方法呈现出以下性质特征。

(1)非线性。一个人工神经元可以是线性或非线性的。一个由非线性神经元互联而成的神经网络自身是非线性的,并且非线性是一种分布于整个网络中的特殊性质。非线性是一个很重要的性质,特别当产生输入信号(如语音信号)内部的物理机制是天生非线性时。

(2)输入输出映射。有监督学习或有教师学习是一个神经网络学习的流行范例,该学习方法涉及使用带标号的训练样本或任务例子对神经网络的突触权值进行修改。每个样本由一个唯一的输入信号和相应期望响应组成。从一个训练集中随机选取一个样本给网络,网络就调整它的突触权值(自由参数),以最小化期望响应和由输入信号以适当的统计准则产生的实际响应之间的差别。使用训练集中的众多样本对该神经网络进行训练,最终到网络没有显著的突触权值修正的稳定状态为止。训练过程中先前用过的例子可能还要在训练期间以不同顺序重复使用。因此,对当前问题网络时通过建立输入/输出映射从例子中进行学习。比如,考虑一个模式分类任务,这里的要求是把代表具体物体或事件的输入信号分类到几个预先分好的类中去。在这个问题的非参数方法中,要求利用例子集估计输入信号空间中模式分类任务的任意判决边界,并且不使用概率分布模型。有监督学习范例隐含了一个类似的观点,这提示神经网络的输入/输出映射和非参数统计推断之间的一个相近的类比。

(3)适应性。神经网络具备适应性是因为网络具有可调整的突触权值,以适应外界环境的变化。对于一个在特定运行环境下训练生成的神经网络,对环境条件不大的变化容易进行重新训练。而且,当它在一个时变环境(即它的统计特性随时间变化)中运行时,网络突触权值就可以设计成随时间变化。用于模式识别、信号处理和控制的神经网络与它的自适应能力耦合,就可以变成能进行自适应模式识别、自适应信号处理和自适应控制的有效工具。作为一个一般规则,在保证系统保持稳定时一个系统的自适应性越好,当要求在一个时变环境下运行时,它的性能就越具鲁棒性。但是,需要强调的是,自适应性不一定导致鲁棒性,实际可能相反。比如,一个暂态自适应系统可能变化过快,以致对寄生干扰有反应,这将引起系统性能的急剧恶化。为最大限度实现自适应性,系统的主要时间常数应该长到可以忽略寄生干扰,而短到可以反映环境的重要变化。这是一个稳定性—可塑性困扰。

(4)证据响应。在模式识别的问题中,神经网络可以设计成既提供不限于选择哪一个特定模式的信息,也提供决策的置信度的信息。后者可以对过于模糊的模式进行推断。有这些信息,网络的分类性能就会改善。

(5)背景的信息。神经网络的特定结构和激活状态代表知识,而网络中每一个神经元都受到其他神经元的影响。

(6)容错性。一个以硬件形式实现后的神经网络有天生容错的潜质,或者鲁棒性计算的能力,意即它的性能在不利运行条件下逐渐下降。比如,一个神经元或它的连接损坏了,存储模式的回忆在质量上被削弱。但是,由于网络信息存储的分布特性,在网络的总体响应严重恶

化之前这种损坏是分散的。因此,原则上,一个神经网络的性能显示了一个缓慢恶化而不是灾难性的失败。有一些关于鲁棒性计算的实验证据,但通常它是不可控的。为了确保网络事实上的容错性,有必要在设计训练网络的算法时采用正确的度量。

(7) VLSI 实现。神经网络的大规模并行性使它具有快速处理某些任务的潜在能力。这一特性使得神经网络很适合用超大规模集成(Very Large Scale Integrated, VLSI)技术实现。VLSI 的一个特殊优点是提供一个以高度分层的方式捕捉真实复杂性行为的方法。

(8) 分析和设计的一致性。基本上,神经网络作为信息处理器具有通用性。这种特征以不同的方式表现出来:

①神经网络中任何神经元结构相同。

②这种共性使得在不同应用中的神经网络共享相同的理论和学习算法成为可能。

③模块化网络可以用模块的无缝集成来实现。

(9) 神经生物类比。神经网络的设计是由对人脑的类比引发的,人脑是一个容错的并行处理,神经生物学家将(人工)神经网络看作是一个解释神经生物现象的研究工具。

2.2 人工神经元模型及表示方法

2.2.1 神经元模型

1. 神经元模型

神经元是神经网络操作的基本信息处理单位,神经元的非线性模型如图 2.1 所示,它是(人工)神经网络的设计基础。一般情况下,我们可将神经元模型分解为 3 个基本元素,分别介绍如下:

(1) 突触或连接。每个神经元都有其突触权值特征。注意突触权值的下标的写法很重要。第一个下标指查询神经元,第二个下标指权值所在的突触的输入端。和人脑的突触不一样,人工神经元的突触权值有一个范围,可以取正值,也可以取负值。

(2) 加法器。用于求输入信号被神经元的相应突触加权的和。这个操作构成一个线性组合器,表示形式如式(2-1)这个求和子式。

(3) 激活函数。用来限制神经元输出振幅。激活函数也称为抑制函数,通过该函数,输出信号被抑制或限制在一个允许的范围内。通常,一个神经元输出的正常幅度范围可写成单位闭区间 $[0, 1]$ 或另一种区间 $[-1, 1]$ 。

图 2.1 的神经元模型也包括一个外

部偏置,记为 b_k 。偏置的作用是根据其为正或负,相应地增加或降低激活函数的网络输入。

对于任一神经元我们可以用一对方程来进行描述:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (2-1)$$

$$y_k = \varphi(u_k + b_k) \quad (2-2)$$

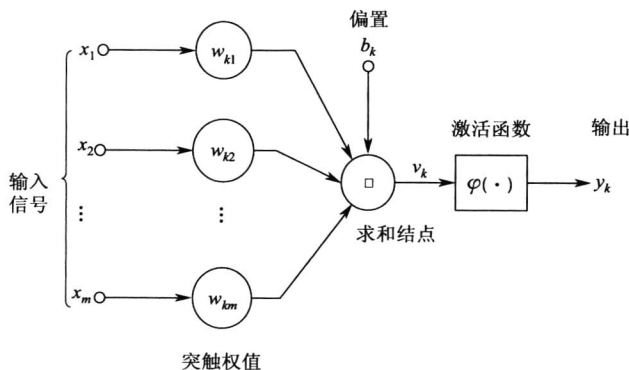


图 2.1 神经元的非线性模型

其中 $x_1, x_2, \dots, x_m$ 是输入信号; $w_{k1}, w_{k2}, \dots, w_{km}$ 是神经元 k 的突触权值; u_k 是输入信号的线性组合器的输出; 偏置是 b_k ; 激活函数是 $\varphi(\cdot)$; y_k 是神经元输出信号。偏置 b_k 的作用是对图2-1中的线性组合器的输出 u_k 进行变换, 变换如下所示:

$$v_k = u_k + b_k \quad (2-3)$$

特别地, 根据偏置 b_k 取正或取负, 神经元 k 的诱导局部域或激活电位 v_k 和线性组合器输出 u_k 的关系如图 2.2 所示。以后我们将使用“诱导局部域”这个术语。由于仿射变换的作用, v_k 与 u_k 的图形不再经过原点。

偏置 b_k 是人工神经元的外部参数。我们可以结合方程(2-1)和(2-3)得到如下公式:

$$v_k = \sum_{j=0}^m w_{kj} x_j \quad (2-4)$$

$$y_k = \varphi(v_k) \quad (2-5)$$

在式(2-4)上加一个新的突触, 其输入是

$$x_0 = +1 \quad (2-6)$$

权值是

$$w_{k0} = b_k \quad (2-7)$$

我们因此得到了神经元 k 的另一个非线性模型如图 2.3 所示。在这个图中, 偏置的作用是做两件事: ①添加新的固定输入 $+1$; ②添加新的等于偏置 b_k 的突触权值。虽然形式上图 2.1 和图 2.3 的模型不相同, 但在数学上它们是等价的。

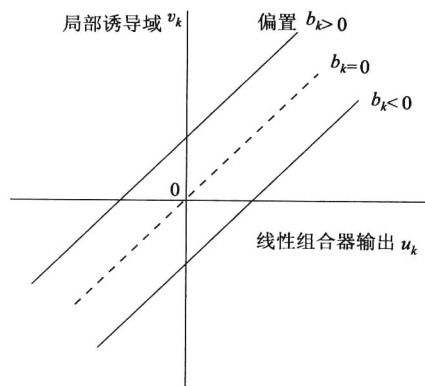


图 2.2 偏置产生的仿射变换(注意 $u_k = 0$ 时, $v_k = b_k$)

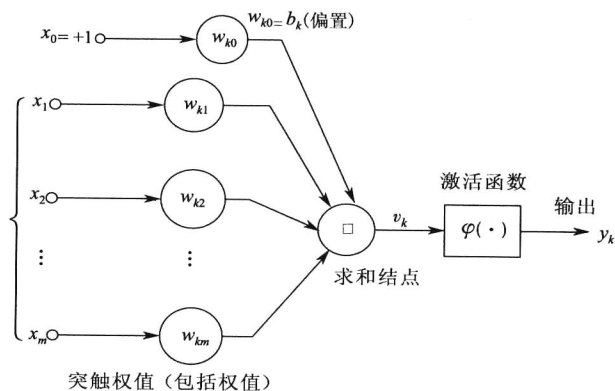


图 2.3 神经元 k 的另一个非线性模型

2. 激活函数的类型

激活函数, 记为 $\varphi(v)$, 通过诱导局部域 $w_{ij} = \begin{cases} -2^{i+j}, & i \neq j \\ 0, & i = j \end{cases}$ 定义神经元输出。这里我们给出 3 种基本的激活函数。

(1) 阈值函数。这种激活函数如图 2.4(a) 所示, 可写为

$$\varphi(v) = \begin{cases} 1 & \text{如果 } v \geq 0 \\ 0 & \text{如果 } v < 0 \end{cases} \quad (2-8)$$

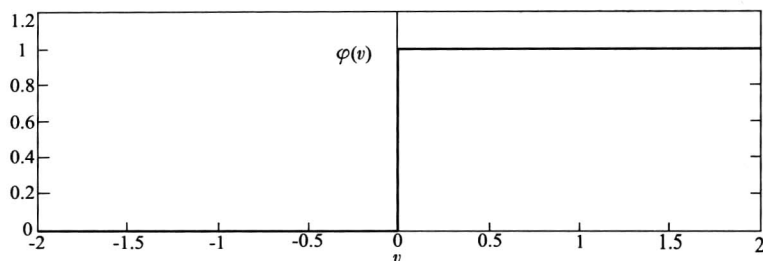
相应地,在神经元 k 使用这种阈值函数,其输出可表示为

$$y_k = \begin{cases} 1 & \text{如果 } v_k \geq 0 \\ 0 & \text{如果 } v_k < 0 \end{cases} \quad (2-9)$$

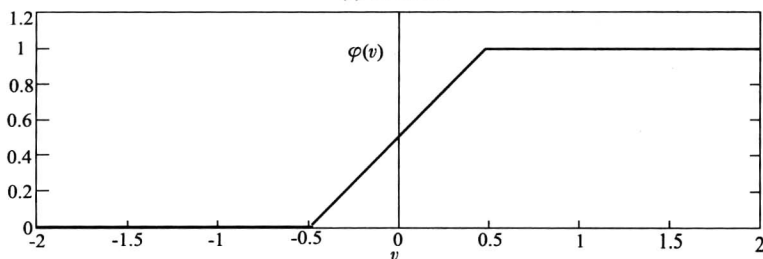
式中, v_k 为神经元的诱导局部域,即

$$v_k = \sum_{j=1}^m w_{kj} x_j + b_k \quad (2-10)$$

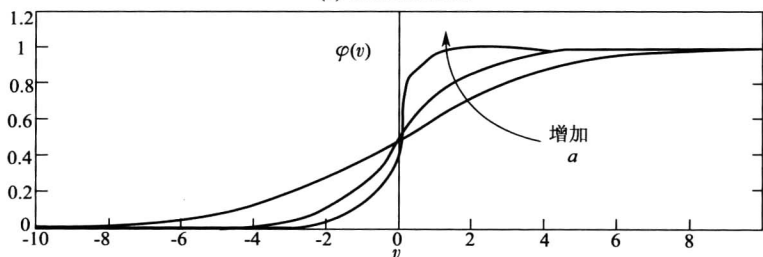
这种模型称为 McCulloch-Pitts 模型,简称 MP 模型,该模型中,如果神经元的诱导局部域非负,则输出为 1,否则为 0。



(a) 阈值函数



(b) 分段线性函数



(c) 具有不同倾斜参数 a 的 Sigmoid 函数

图 2.4 3 种基本的激活函数

(2) 分段函数。分段线性函数如图 2.4(b) 所示,则有:

$$\varphi(v) = \begin{cases} 1 & v \geq +\frac{1}{2} \\ v & +\frac{1}{2} > v > -\frac{1}{2} \\ 0 & v \leq -\frac{1}{2} \end{cases} \quad (2-11)$$

式中,在运算的线性区域内放大因子置为 1。这种形式的激活函数是对非线性放大器的近似。