

人工神经网络原理与应用

(第二版)

王 旭 王 宏 王文辉 编著

东北大学出版社

· 沈 阳 ·

© 王 旭 王 宏 王文辉 2007

图书在版编目 (CIP) 数据

人工神经网络原理与应用 / 王旭, 王宏, 王文辉编著. —2 版. —沈阳: 东北大学出版社, 2007.8

ISBN 978-7-81102-451-7

I. 人... II. ①王... ②王... ③王... III. 人工神经网络—高等学校—教材
IV. TP183

中国版本图书馆 CIP 数据核字 (2007) 第 129830 号

内 容 提 要

本书以通俗易懂的方式讲述了人工神经网络的基本原理、设计和计算方法。全书共分十二章。第一章介绍了人工神经网络的发展历史和它的主要特点; 第二章和第三章介绍了生物神经网络和人工神经网络的基本原理; 第四章至第十章讲述了几种主要的人工神经网络的原理和具体计算方法; 第十章介绍了人工神经网络在自动控制系统中的应用; 第十章介绍了人工神经网络在自动控制系统中的应用; 最后一章介绍了编程实例; 全书配有习题。

本书可作为自动控制、计算机应用、通信工程等有关专业大学本科及研究生的教材, 也可供相关领域的工程技术人员和研究人员自学和参考。

出 版 者: 东北大学出版社

地址: 沈阳市和平区文化路 3 号巷 11 号

邮编: 110004

电话: 024—83687331 (市场部) 83680267 (社务室)

传真: 024—83680180 (市场部) 83680265 (社务室)

E-mail: neuph@neupress.com

http://www.neupress.com

印 刷 者: 沈阳市政二公司印刷厂

发 行 者: 东北大学出版社

幅面尺寸: 184mm×260mm

印 张: 10

字 数: 250 千字

出版时间: 2007 年 8 月第 2 版

印刷时间: 2007 年 8 月第 1 次印刷

印 数: 2000 册

责任编辑: 李毓兴

封面设计: 唐敏智

责任校对: 薛 升

责任出版: 杨华宁

ISBN 978-7-81102-451-7

定 价: 20.00 元

第二版前言

第一版问世至今，本学科又有了长足发展，为了反映最新的科技成果和便于读者学习，本版改正了第一版中的差错，调整了部分内容，重写了第十二章。

尽管如此，书中仍难免有不足之处，企盼读者批评指正。

编者

2007年8月

第一版前言

随着科学技术的高速发展，人们对自动控制和优化计算等方面的要求越来越高，迫切需要提高对信息处理的智能化水平。由于人工神经网络是模仿人的脑神经功能而提出来的，它具有很强的自学习能力，能适应于复杂环境和多目标控制的要求，所以引起了人们广泛的关注，并在许多领域中起到了很大的作用。

本书作为人工神经网络方面的入门书，由浅入深详细地介绍了人工神经网络的原理与特点，对于每一种常用的网络都给出了详细的计算方法。本书共分十二章。第一章介绍了人工神经网络的发展史，并以一个例子讲述了它的主要特点；第二章介绍了生物神经网络的基本原理，可使读者对于真实的神经网络有一些基本的概念，这对于理解和研究人工神经网络是必不可少的；第三章叙述了人工神经网络的结构和分类；第四章至第十一章讲述了几种主要的人工神经网络的原理和具体计算方法；最后一章介绍了一些编程实例。

本书第一章至第五章由王宏编写；第六章至第十二章由王旭编写，第一、二、七章习题由王旭编写；第三、四、五、六、八、九、十和十一章的习题由王文辉编写。全书由王旭负责统稿。在编写过程中，参阅了国内外大量的文献和资料，对此编者表示衷心地感谢。

由于编者水平有限，加之人工神经网络理论本身还处于发展阶段，书中如有不妥之处，恳切欢迎读者批评指正。

编 者

2000年12月

目 录

第一章 绪 论.....	1
1.1 为什么要用人工神经网络	1
1.2 人工神经网络的发展	3
1.3 人工神经网络是怎样工作的	4
习 题	12
第二章 生物神经网络的基本原理	13
2.1 生物神经元.....	13
2.2 生物神经网络的结构.....	15
习 题	18
第三章 人工神经网络的基本原理	19
3.1 人工神经元.....	19
3.2 人工神经网络模型.....	22
3.3 神经网络的学习过程.....	28
3.4 神经网络的学习规则.....	30
3.5 神经网络的工作过程.....	37
习 题	38
第四章 感知机模型网络	43
4.1 感知机网络的结构.....	43
4.2 感知机网络的学习规则.....	43
4.3 感知机网络的局限性.....	47
习 题	49
第五章 多层网络的误差逆传播校正方法	51
5.1 误差逆传播校正方法.....	51
5.2 BP 网络的学习规则与计算方法	54
5.3 BP 网络的应用举例	59
5.4 BP 网络的改进方案	63
习 题	65
第六章 Hopfield 神经网络	69
6.1 Hopfield 网络的基本结构	69
6.2 Hopfield 网络的能量函数与运行规则	71
6.3 联想记忆.....	74

6.4	Hopfield 网络连接权的设计方法	75
6.5	Hopfield 网络的弱点	78
6.6	连续型 Hopfield 神经网络.....	79
6.7	Hopfield 网络的应用	82
	习 题	86
第七章	随机型神经元网络	93
7.1	模拟退火算法.....	93
7.2	波尔茨曼机模型及其工作规则.....	95
7.3	波尔茨曼机模型网络的学习规则.....	98
	习 题.....	103
第八章	竞争型神经元网络.....	104
8.1	竞争型神经网络的基本结构及其学习方法	104
8.2	抑制竞争型神经网络及其学习规则	107
8.3	自适应共振理论神经网络	109
	习 题.....	114
第九章	自组织特征映射神经网络.....	116
9.1	SOFM 网络结构及学习工作规则	116
9.2	SOFM 网络的应用	120
	习 题.....	122
第十章	对向传播神经网络.....	124
10.1	CP 网络的结构及学习工作规则	124
10.2	CP 网络的改进与完善	126
	习 题.....	128
第十一章	神经网络控制简介.....	130
11.1	神经网络控制的基本原理.....	130
11.2	几种典型的神经网络控制系统的结构.....	131
	习 题.....	133
第十二章	人工神经网络编程实例.....	135
12.1	感知机模型神经网络编程.....	135
12.2	BP 神经网络编程	139
12.3	ART1 型神经网络编程	146
	参考文献.....	150

第一章 绪 论

1.1 为什么要用人工神经网络

自从 20 世纪 80 年代初以来,人工神经网络技术就在全世界范围内迅速地发展起来。其中的主要原因是人工神经网络的结构和它的计算方法是在模拟人类大脑的结构和思维方式的基础上建立起来的。

人的大脑实际上是由很复杂的神经网络所组成的,正是因为这些神经网络的作用,人才能以很高的速度理解感觉器官传来的信息,尽管这些信息有时含糊不清。比如:人的听觉神经网络能够在喧闹的环境中识别清楚对方的声音;人的视觉神经网络能够在不到 1 秒钟的瞬间认出多年未见的老朋友;人的智能神经网络能够归纳出某一长篇文章的中心思想;另外还有一个最重要的方面,就是人脑的学习能力和创造能力,它能从环境中学习,从书本中学习,从所经历的事件中学习,并能利用所学到的知识去创造新的知识。这所有的一切都是目前任何一种人造的机器所望尘莫及的。人脑之所以有这些能力,主要取决于人脑本身的构造,加之多年以来的学习过程。因此,为了解释人脑的奥秘,也为了提高目前人造机器的能力,人们对人脑的工作机理进行了长期不懈的探索和模拟,逐渐发展起来了一门新兴的学科——人工神经网络。这一学科的研究目的在于模拟大脑神经网络的原理来构成一种人造的“大脑”,或者说将目前的电脑变成“大脑”。

到目前为止,人工神经网络系统已经具有一些同人脑相类似的特点,在信息的分布式存储、数据的并行处理以及利用外来的信息进行自学习的能力等方面都同人脑很相似。所以它在智能控制、信息处理、模式识别等领域已经起着重要的作用并有着广阔的应用前景。今天,随着科学技术的迅猛发展,对神经网络研究的重要意义不亚于对数字计算机的研究。

与人脑相比,数字计算机有许多不及人脑的方面。首先人脑有很强的自学习能力。例如,人刚出生时除了本能的动作之外,大脑中并没有存储多少信息,也没有什么知识,但是随着时间的推移,通过对外界环境的感知以及读书学习的过程,人脑中存储的信息会逐渐增多,变得越来越聪明。这种学习的能力是人脑自身的功能,是自然而然的事情。在这一方面数字计算机就显得那样的机械和无能。

人脑的联想记忆能力也是很卓越的。人脑在存贮大量的信息(包括图像、语言、文字、形状等)之后,能对这些信息进行综合分析处理,当有感官信号进入时,脑神经系统能够自行组织,很快地联想到一系列相关的存贮信息,并迅速做出反应。例如,见到一位老朋友时,人脑就能联想到老朋友的名字、爱好与特长,或许还能想起曾经共同度过的一段美好时光。

人脑具有很强的识别能力。它能够对变形的、模糊的、残缺的信息进行快速而正确的识别,复原出原来记忆的形状。在这方面传统的计算机就显得无能为力。

高稳定性。尽管每天大脑中都有神经细胞的正常死亡,但这并不影响大脑的功能。大

脑的局部损伤也只会引起某些脑功能的减弱,而不会使功能突然全部丧失,这与传统的数字计算机是不同的。对一般的数字计算机来讲,不同的数据信息是按地址存放在存储器中,它们是互不相关的。仅当程序运行时某些信息才能相互关联。这种联系要通过复杂的算法才能实现,而且受运行程序的限制很大。软件、硬件上的任何微小差错都可能导致严重的错误结果。所以,传统的数字计算机的稳定性比人脑要差得多。

人脑具有很强的分析能力和判断能力。人脑善于利用获取的信息和学到的知识进行综合分析,并利用分析的结果来处理日常遇到的事件,有时甚至能产生出创造性的想法。例如,早晨出门前看到天气阴得很,这时人们会考虑是带雨伞还是不带。若带雨伞,天不下雨岂不累赘?若不带雨伞,一旦下雨挨淋怎么办?会不会感冒?等等。对此事的判断,反映了人的分析判断能力。分析判断不仅仅是一个简单的逻辑推理过程,它与人已有的经验以及许多因素有关。人脑所具有的这种分析、判断能力以及创造能力是数字计算机无法达到的。

人脑有大规模并行处理能力。人的大脑中有上百亿个神经细胞,它们可以同时起作用,并行处理信息,并利用处理的结果作出判断。所以它的工作速度要比传统的计算机快许多倍。

由于人脑的神经网络有以上这么多非凡的能力,所以人们想模拟人脑神经网络的原理和结构来研究出人工神经网络,以解决目前科学技术上日趋复杂的难题。

实际上,对人工神经网络研究的另一个重要目的是要创造出另一类的计算机,一种具有人脑功能的新型计算机——神经计算机。神经计算机是基于人们对人脑的组织结构、活动机理的认识而提出来的一类全新的计算机。它与目前的数字计算机相比有许多本质的区别,它能模仿人脑神经系统的组织结构和活动机理,可以通过学习获得对外界事物的认识;它以分布式存储的方式存储数据、积累知识。当给它以很少的提示信息时,便能很快地回忆出其整体信息和相关信息。

神经计算机的硬件部分是基于人工神经网络的原理,利用现有的微电子技术以及正在发展的分子生物器件来实现的。它采用集成电路芯片模拟脑神经细胞,并将这些细胞元件按一定的结构连接起来,构成一个并行分布式的网络系统。神经计算机具有很强的对外界环境的适应能力。当外界环境发生变化时,经过一段时间的训练或感知,神经网络能自动调整网络的结构和参数,使它始终处于一种最佳的工作状态。

神经计算机是由大量的简单处理元件构成的一个大系统,在信息处理时依据神经网络的原理,具有大规模并行处理的特征。因此,也就具有很强的容错能力,部分信息的丢失或模糊信息的输入仍可以得到完整的结果,表现出很强的错容性。

神经计算机与现有的数字计算机相比,它们有各自的适用场合。数字计算机善于处理结构化问题,在逻辑推理、数值计算方面表现出很强的能力,并非人脑可比拟的。神经计算机则善于处理形象思维领域非结构化的问题,二者的处理能力具有互补性。其次,处理问题的方法不同。传统计算机是基于算法、顺序执行程序求解的;神经计算机不用事先编写程序,它是通过学习、自适应来求解问题的,采用大规模并行处理。因此,神经计算机与数字计算机将在各自不同的领域中为科学技术的发展起到重要的作用。

目前,自动控制领域面临的主要难题是控制对象越来越复杂,有多种不确定的、难以确切描述的非线性问题。这样就对控制系统的要求越来越高,要求智能化的控制方法。由于人工神经网络源于对脑神经的模拟,有很强的适应于复杂环境的能力和对多目标控制有很

好的自学习能力,并可以任意精度逼近任意非线性连续函数的特性。这正好能解决控制领域的许多难题。

1.2 人工神经网络的发展

人们对神经网络的研究是从 20 世纪 40 年代初就开始了,至今已有半个多世纪的历史。1943 年,心理学家 M. McCulloch 和数学家 W. H. Pitts 采用数理模型的方法首先提出了一种人工神经元模型,简称 M—P 模型,迈出了人类研究神经网络的第一步。1949 年心理学家 D. O. Hebb 在他所著的《行为组织学》(Organization of Behavior)一书中提出了神经元之间连接强度的变化规律,即:当细胞 A 的轴突传导的神经信号刺激细胞 B 时,如果细胞 B 因此而产生动作电位,则 AB 间的突触连接加强;如果细胞 B 不产生动作电位,则 AB 间的连接将减弱。这就是我们所说的 Hebb 学习规则。这一规则在当时是作为假设提出来的,其正确性在 30 年后得到了证实,至今它在各种神经网络模型研究中仍起着重要作用。

第一个智能型的人工神经网络系统是在 1958 年由 F. Rosenblatt 提出的感知机(Perceptron)模型网络。在这种人工神经网络中,记忆的信息存贮在连接权上,外部刺激通过连接通道自动激活相应的神经元,以达到自动识别的目的。虽然感知机模型比较简单,但已经具有神经网络的一般性质,如可学习性、并行处理、分布式存贮等,这些性质与当时流行的串行、离散的符号处理的电子计算机与人工智能技术完全不同,因此,引起了研究人员的极大兴趣。而且这种网络一直到现在还在广泛应用。1960 年,B. Widrow 和 M. Hoff 从工程角度出发发表了“自适应开关电路”一文,提出了一种自适应线性神经元模型(Adaline)以及一种有效的网络学习方法,即通常所说的 Widrow—Hoff 学习规则(或称 δ 学习规则)。而且他们对神经网络硬件系统的开发也有相当深入的研究。

随着人工神经网络研究的不断深入,人们曾遇到过来自各方面的各种困难和许多一时难以解决的问题。对神经网络的学习能力问题,引起了学术界的很大争议。人工智能的创始人之一——M. Minsky 和 S. Papert 于 1969 年发表了对神经网络研究产生重要影响的《感知机》(Perceptron)一书。书中提出了感知机网络的局限性,它大大地影响了人们对神经网络研究的兴趣。使人工神经网络的研究在 70 年代处于低潮。

但是,在神经网络研究处于低潮的这一时期,仍有为数不多的学者不遗余力地致力于神经网络的研究。1969 年,美国学者 S. Grossberg 与 G. A. Carpenter 提出了自适应共振理论(ART)模型,并在以后的若干年发展了 ART1, ART2, ART3 三种神经网络模型。1972 年,芬兰学者 T. Kohonen 提出了自组织映射(SOM)理论网络。SOM 模型是一类很重要的无监督学习型人工神经网络,主要用于模式识别、语音识别、数据分类等方面。

在人工神经网络的研究被冷落了 10 年之后的 1982 年,美国物理学家 Hopfield 教授发表的一篇突破性的学术论文,他提出了一种新的神经网络模型,并指出可以用微电子器件来实现它,这个神经网络是基于磁场的结构特征而提出来的,很容易被工程技术人员理解,引起了人们的普遍关注。后来人们称它为 Hopfield 神经网络。在其网络模型中,定义了神经网络的“能量函数”,并给出了网络稳定性的判据,使所提出的网络具有联想记忆和对优化问题求解的能力。特别是 Hopfield 将这种网络成功地运用于著名的“巡回推销商”(TSP)问题的求解,取得了很好的效果。Hopfield 神经网络的出现为神经计算机

(Neurocomputer)的研制奠定了基础,并开创了神经网络用于联想记忆和优化计算的新途径。同时也使被冷落 10 年之久的对人工神经网络的研究复兴起来了。

在 1984 年之后的三年内,美国 AT&T 公司贝尔实验室宣布了利用 Hopfield 网络原理实现了第一个基于硅芯片的硬件神经网络。G. E. Hinton 和 T. J. Sejnowski 借助统计物理学的概念和方法提出了一种随机神经网络模型——波尔茨曼机,学习过程中采用模拟退火原理,有效地克服了 Hopfield 网络存在的能量局部极小问题,使整个网络系统最终能达到能量全局最小点,也就是可以获得一个最优解。

1986 年至 1988 年 D. E. Rumelhart 与 L. L. McClelland 领导的 PDP 研究小组出版了《并行分布处理》一书,全面介绍了 PDP 理论。他们将神经网络模型归结为具有三个基本属性:结构、神经节点传递函数及学习算法。并发展了适用于多层网络学习的误差逆传播(BP)算法,为解决多层网络学习难的问题开辟了一条道路。BP 算法是迄今为止应用最普遍的神经网络学习算法。由于在 20 世纪 80 年代大量而深入的开拓性工作的出现,大大发展了人工神经网络的理论,使人们对模仿大脑来进行信息处理充满了希望。

近几年来,神经网络理论的研究与实现引起了美国、日本、中国及西欧一些国家的科学家、研究机构和企业界的普遍关注。并且各个学科的研究人员都想利用人工神经网络的特殊功能来解决本学科的难题,很多的工程项目都采用和正准备采用人工神经网络的解决方案。同时不同学科的科学工作者正在积极地联合起来进行各种学术交流。可以说,目前对人工神经网络的研究出现了更高的热潮。

目前,各种有关人工神经网络的专著逐年增加,许多期刊上不断推出研究人工神经网络的专集。在世界范围内已经形成了研究神经网络的前所未有的热潮。可以预言,21 世纪初人工神经网络理论将会有更大的发展,它的应用将推动科学技术的大步前进。

1.3 人工神经网络是怎样工作的

为了让大家对人工神经网络有一个具体的印象,下面我们以一个例子来描述一下人工神经网络的工作方式。比如说在天气预报中有这样一段广播:“……今天下午将有一股强冷空气袭击沈阳上空,受它的影响今天下午到夜间我市将有中到大雨……”。

这时可能将出现这样一段情景:首先看到窗外一片黑云,随后气温下降、下起雨来,由于天气变坏,去露天游泳的计划就得取消了,想去游泳的人的心情就会不好。

对于这个例子,我们可以用人工神经网络技术来讨论天气与人们行动和情绪的关系。

图 1-1 示出了由两层组成的一个模型。每层有三个神经元。底层叫输入层,顶层叫输出层。输入层的三个神经元从外界获得信息,而输出层的三个神经元将模型中的信息向外界发出。输入层的每个神经元与输出层的每个神经元相连。

这个模型的信息处理过程是从下层向上层进行的。即输入层得到了天气信息,对此,输入层的每个神经元可以处于不同的状态,即不兴奋状态(天气不好,或天气不变化,或不下雨)或者处于兴奋状态(天气好,或天气变化,或下雨)。输出层发出模型处理的结果。输出层的每个神经元也可以处于不同的状态,即不兴奋状态(不能游泳,或情绪不好,或没带雨伞)或者处于兴奋状态(可以游泳,或情绪好,或带雨伞),图 1-2 示出了一种输入与输出均处于兴奋状态的框图。

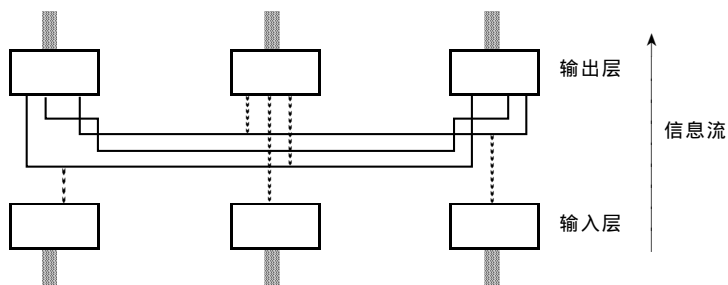


图 1-1 两层结构模型

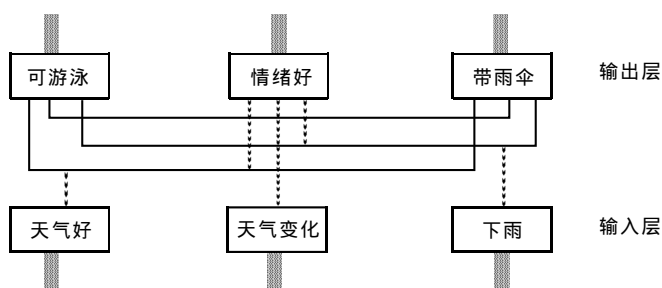


图 1-2 输入与输出均处于兴奋状态的框图

图 1-3 示出了一种期望的输入和期望的输出模型，即“天气不好——不能游泳”；“天气不变化——情绪不好”；“下雨——带雨伞”。我们看一下这个模型是怎样工作的。

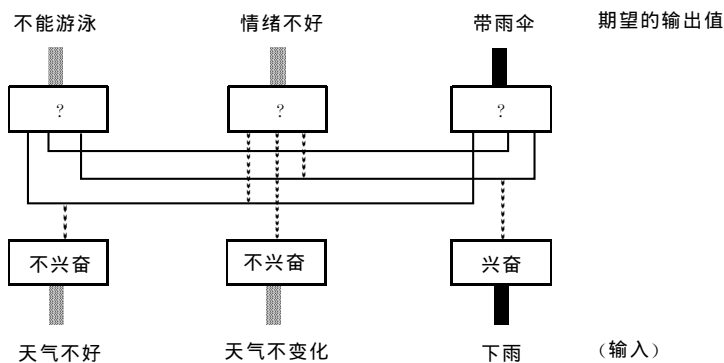


图 1-3 一种期望输入和期望输出模型

图中的输入是“下雨”，得到的输出结果应是“带雨伞”（即情绪不好和不能游泳）。那么首先输入和输出就确定了。为了计算，我们定义“兴奋状态”用 +1 代表；“不兴奋状态”用 -1 代表。这样一来，图 1-3 可表示为图 1-4 的形式。

在这个模型中，每个神经元最少有一个输入信息。对于输入层的神经元来说，信息是从外界进入的。而输出层的神经元中的信息来自下层神经元。通常一个神经元可能有多个输入突触，但只有一个输出端。这个输出值是对所有输入值处理后的结果。而对输入信息进行处理的第一步是对输入值的求和，如图 1-5 所示。

由图 1-5 可见，对输入结果简单地求和是不能得到期望的输出结果的（图中输入信息为下雨，但输出竟是没带雨伞）。那末为了得到理想的结果，必须对每个神经元的输入值进行特殊地处理——加权处理。每个神经元都有不同的加权值，加权处理在实际应用中有很重要的意义。比如可能有这样一个人（这里简称为 K），虽然此刻在下雨，但因为几个小时以后

他将乘飞机去南方旅游。所以此时 K 的心情很好,也就是说外出旅游带来的喜悦心情比下雨而产生的失落感影响要大得多,所以对 K 来说外出旅游的加权值要大于下雨的加权值。

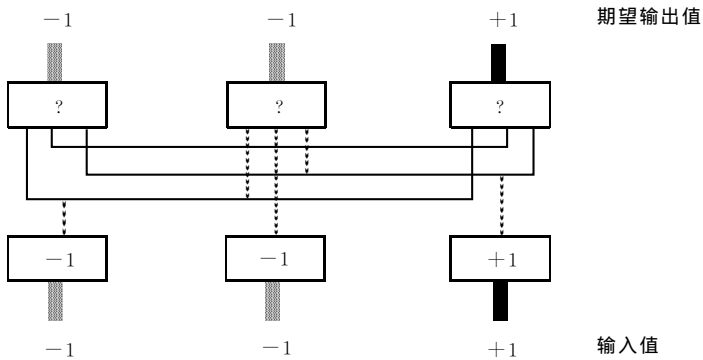


图 1-4 将图 1-3 中的“兴奋状态”用+1 代表;“不兴奋状态”用-1 代表

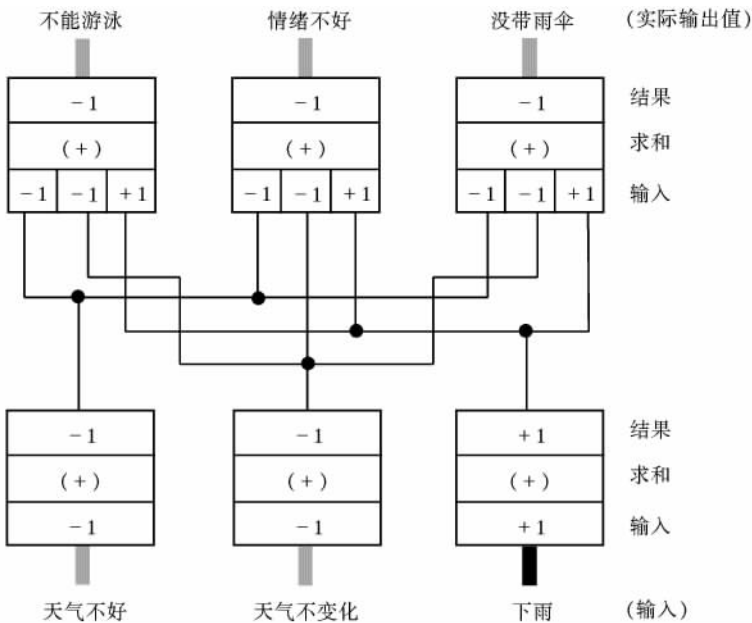


图 1-5 对输入值进行求和处理

考虑加权后的处理过程如图 1-6 所示,在这个模型中,对输入信息先进行加权后再做求和处理。因为我们不能预测实际处理的结果是否如同假设的一样,所以在处理中还要引入一个阈值,如结果大于阈值输出则为兴奋状态(+1),如结果小于阈值输出为不兴奋状态(-1)(如图 1-7 所示),这里阈值取为 0。这样在输入为“下雨”时,就可以得到理想的输出结果——带雨伞(如图 1-8 所示)。

我们还可以对下面的期望输入和期望输出进行训练,即“天气好——游泳,天气不变化——情绪好,不下雨——没带雨伞”。若只考虑加权值和阈值,这个模型如图 1-9 所示,即对输入为“天气不变化”而输出则是“情绪不好”,可见没有得到期望的输出,所以还应考虑其他的因素。

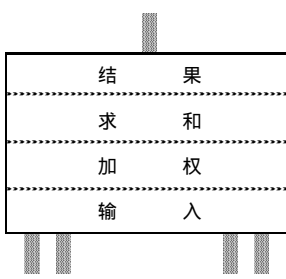


图 1-6 对输入的加权求和过程



图 1-7 加权求和的结果与阈值相比较

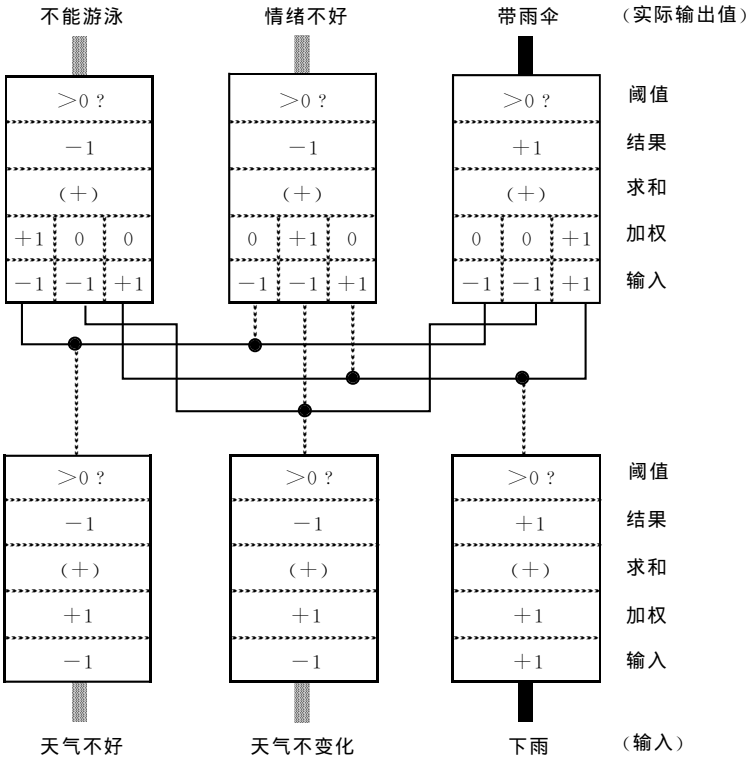


图 1-8 一种期望输入与期望输出的处理过程

因为神经网络是模拟大脑进行工作的，所以它具有学习的能力。这样，可以让神经网络根据输入和期望的输出值进行学习训练。通过学习训练得到合适的加权值。其中的学习规则是根据 Hebb 假说，即：如果神经元 i 和 j 同时兴奋，则神经元 i 和 j 间的连接加强。加强的程度可以用一个叫学习系数的实数来表示。

对于上面的例子，我们也可以根据 Hebb 假说再进行修正。在上例中，输入层左边的神经元和期望输出层左边、中间神经元都是兴奋的。根据 Hebb 学习规则这些神经元之间的连接权的值应加强，比如取学习系数为 1（即提高了 1），结果如图 1-10 所示可见加权以后求和的结果对第一个神经来说是 2，对第二个神经是 0。因这里学习系数取为 1，所以输出层的神经元没有达到期望的输出值。如果重新学习，使学习系数的加权值再提高 1，则可以得到期望的输出值。这样一来，整个模型的连接强度也发生了变化，就会得到图 1-11。

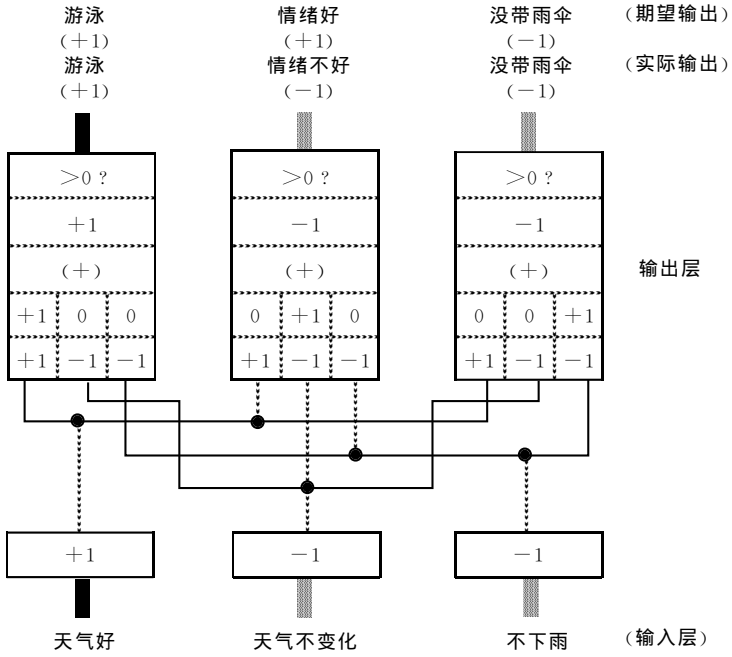


图 1-9 期望输出与实际输出的比较

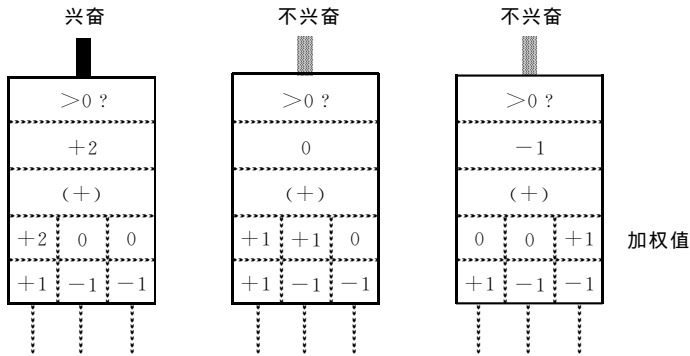


图 1-10 根据 Hebb 假说对连接权修正

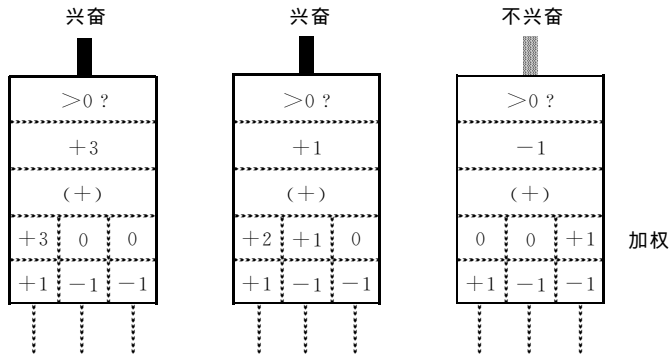


图 1-11 对连接权再一次修正

这样就得到了最后的期望输出结果,如图 1-12 所示。

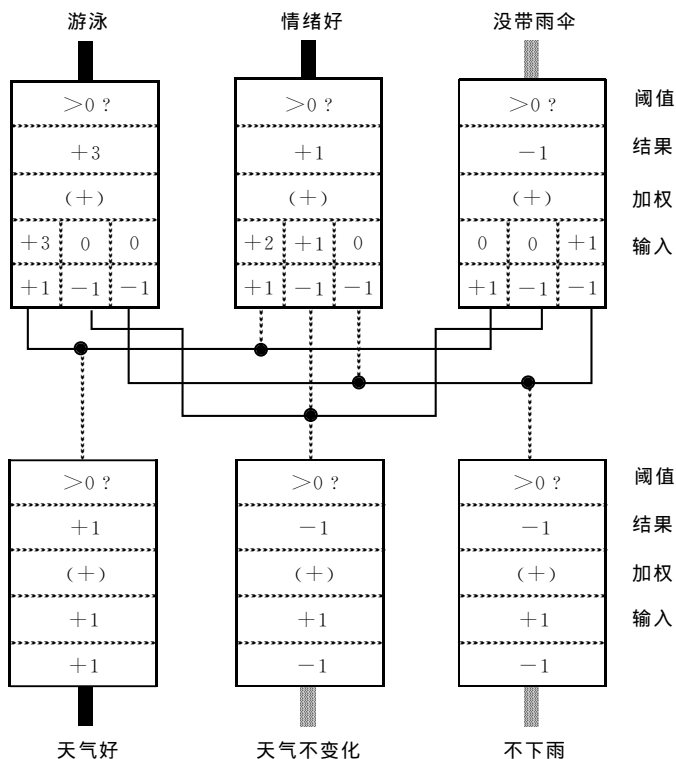


图 1-12 实际输出与期望输出相符

虽然对于这样一个简单学习过程的学习目的很快就达到了。但当三个输入都是兴奋状态,并且三个期望输出都是不兴奋状态时,上述就不适用了。即用目前的连接强度和阈值给出的是三个兴奋的输入引起三个兴奋的输出(如图 1-13 所示)。

由图 1-13 可见,仅使用上述的学习规则得不到期望的输出值,所以必须还要考虑其他因素,即要考虑到神经元的期望输出值与实际输出值的差别。这个差别叫误差。根据这个误差值与学习系数的乘积在图 1-13 的基础上再来确定新的加权值。

误差值 = 期望输出值 - 实际输出值

新的加权值 = 旧的加权值 + (学习系数 × 误差值)

对于上面的例子,令:学习系数 = 0.5

W_{ij} = 神经元 j 到神经元 i 的连接权

在第一次学习中,计算结果如下:

$$W_{出1,入1} = 3 + 0.5 \times (-1 - 1) = +2$$

$$W_{出1,入2} = 0 + 0.5 \times (-1 - 1) = -1$$

$$W_{出1,入3} = 0 + 0.5 \times (-1 - 1) = -1$$

$$W_{出2,入1} = 2 + 0.5 \times (-1 - 1) = +1$$

$$W_{出2,入2} = 1 + 0.5 \times (-1 - 1) = 0$$

$$W_{出2,入3} = 0 + 0.5 \times (-1 - 1) = -1$$

$$W_{出3,入1} = 0 + 0.5 \times (-1 - 1) = -1$$

$$W_{出3,入2} = 0 + 0.5 \times (-1 - 1) = -1$$

$$W_{出3,入3} = 1 + 0.5 \times (-1 - 1) = 0$$

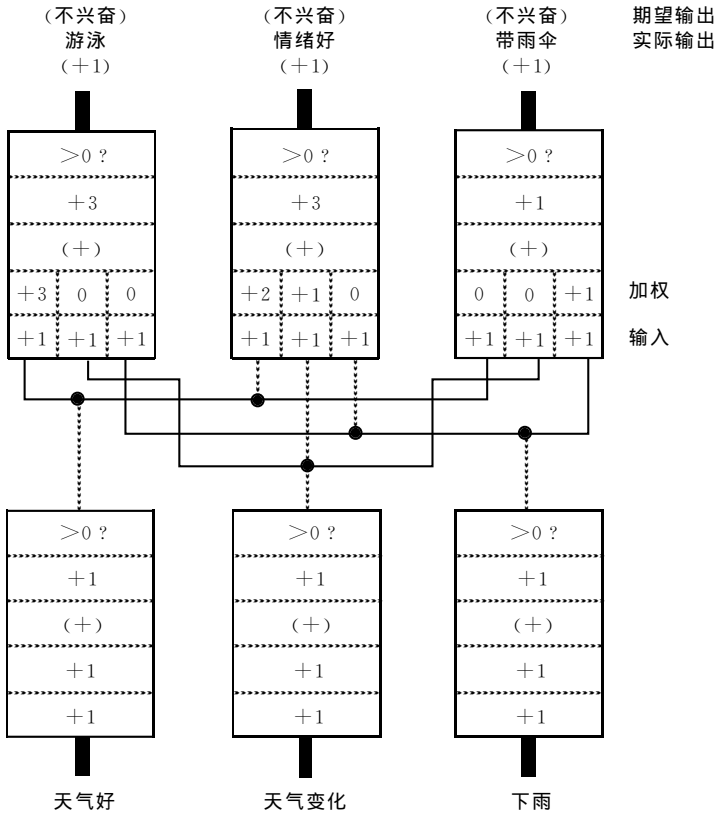


图 1-13 实际输出与期望输出不相符

学习结果如图 1-14 所示,再在图 1-14 的基础上进行第二次学习训练,计算结果如下:

$$\begin{aligned}
 W_{出1,\lambda1} &= 2 + 0.5 \times (-1 + 1) = +2 \\
 W_{出1,\lambda2} &= -1 + 0.5 \times (-1 + 1) = -1 \\
 W_{出1,\lambda3} &= -1 + 0.5 \times (-1 + 1) = -1 \\
 W_{出2,\lambda1} &= +1 + 0.5 \times (-1 + 1) = +1 \\
 W_{出2,\lambda2} &= 0 + 0.5 \times (-1 + 1) = 0 \\
 W_{出2,\lambda3} &= -1 + 0.5 \times (-1 + 1) = -1 \\
 W_{出3,\lambda1} &= -1 + 0.5 \times (-1 + 1) = -1 \\
 W_{出3,\lambda2} &= -1 + 0.5 \times (-1 + 1) = -1 \\
 W_{出3,\lambda3} &= 0 + 0.5 \times (-1 + 1) = 0
 \end{aligned}$$

可见在第二次学习中,结果没有发生变化,即误差是零,这时学习训练结束。这样对于三个输入都是兴奋时的计算结果如图 1-14 所示。这个模型是对三个兴奋输入值进行的训练。最后它学会了识别这个模型并产生期望的输出(三个不兴奋输出)。根据这样的学习训练取加权值,可以用在图 1-15 的例子中。

可见“带雨伞的例子”已正确识别了,那么“天气好的例子”也将正确地处理(如图1-16所示)。

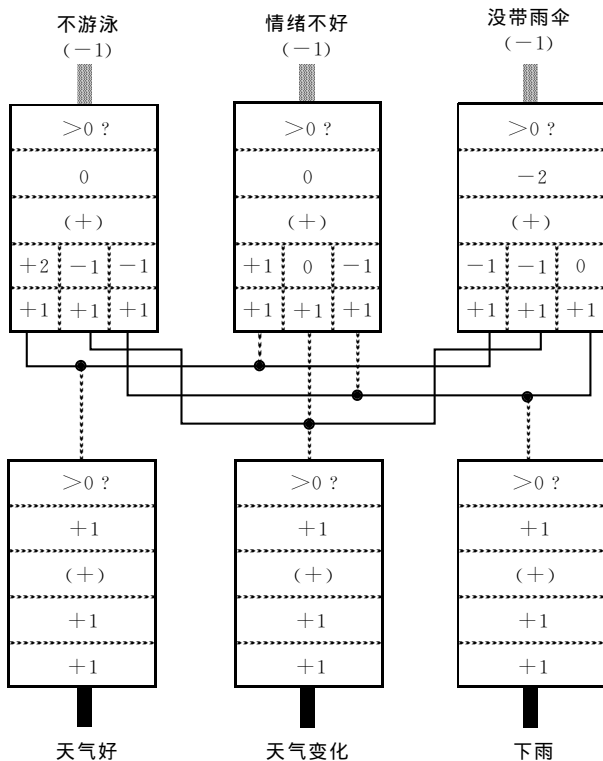


图 1-14 考虑误差因素后的处理结果

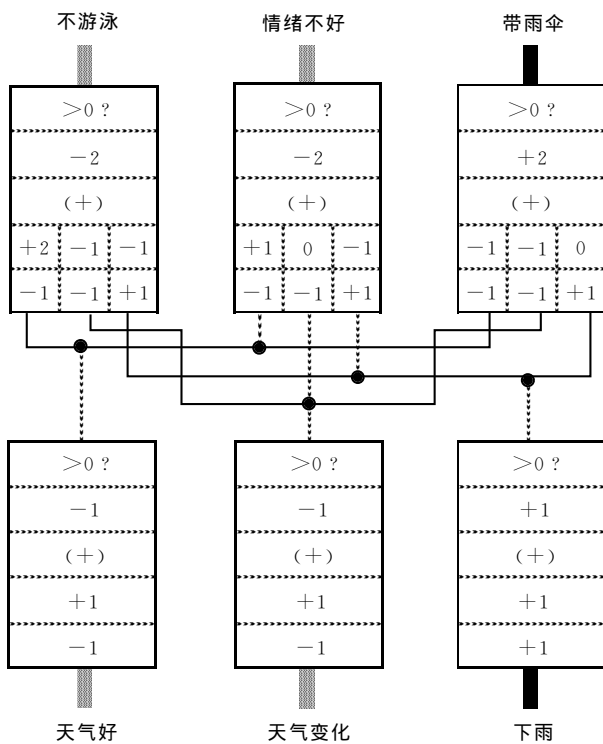


图 1-15 一种期望输出与实际输出相符的情况