

“国家自然科学基金”资助项目

神经网络和遗传算法 在水科学领域的应用

苑希民 李鸿雁 刘树坤 崔广涛 著



中国水利水电出版社
www.waterpub.com.cn

“国家自然科学基金”资助项目

神经网络和遗传算法 在水科学领域的应用

苑希民 李鸿雁 刘树坤 崔广涛 著



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书较全面系统地介绍了应用人工神经网络和遗传算法解决水科学问题的最新研究成果。在理论方面：将人工神经网络技术应用于解决复杂、模糊、高度非线性洪水、泥沙的预测预报问题，提出了基于人工神经网络的峰值识别理论，采用遗传算法优化神经网络的初始权重，实现了人工神经网络与遗传算法的有机结合。在应用方面：研制了基于人工神经网络与遗传算法理论的洪水预报系统和多泥沙洪水预报系统。此外，作者在调查和分析国内外大量研究成果的基础上，综述了人工神经网络技术在水资源预报和优化调度、环境污染预测、卫星和遥感图像的识别等方面的应用。本书结合实例，深入阐述了神经网络技术和遗传算法理论、方法、最新成果在水科学领域的应用，是理论联系实际的经验总结。本书适用于水利、气象、环境、遥感信息领域的技术和科研人员，对人工智能和优化技术应用研究的有关学者、高校师生有重要的参考价值。

图书在版编目 (CIP) 数据

神经网络和遗传算法在水科学领域的应用/苑希民等著. —北京：中国水利水电出版社，2002

ISBN 7-5084-1085-8

I. 神… II. 苑… III. ①神经网络-应用-水利工程-工程技术②遗传-算法-应用-水利工程-工程技术 IV. TV-39

中国版本图书馆 CIP 数据核字 (2002) 第 029048 号

书 名	神经网络和遗传算法在水科学领域的应用
作 者	苑希民 李鸿雁 刘树坤 崔广涛 著
出版、发行	中国水利水电出版社 (北京市三里河路 6 号 100044) 网址: www.waterpub.com.cn E-mail: sale@waterpub.com.cn 电话: (010) 63202266 (总机)、68331835 (发行部)
经 售	全国各地新华书店
排 版	中国水利水电出版社微机排版中心
印 刷	涿州市星河印刷厂
规 格	787×1092 毫米 16 开本 11.5 印张 273 千字
版 次	2002 年 8 月第一版 2002 年 8 月第一次印刷
印 数	0001—3000 册
定 价	48.00 元 (含光盘 1 张)

凡购买我社图书，如有缺页、倒页、脱页的，本社发行部负责调换

版权所有·侵权必究

随着人类社会的发展、文明的进步以及由此带来的“科技以人为本”理念的不断深化,旨在代替人类从事危险、繁重、重复劳动的研究显得极为重要,一些模拟人脑对客观对象进行识别、分析、判断、处理,即对具有复杂、模糊、高度非线性问题识别能力的知识体系研究,引起科学工作者的广泛兴趣。

建立在 Von Neumann 原理基础上的算法和程序存取式传统计算机,在计算速度、计算结果精度和可靠性方面都远远超过了人脑,而对模糊信息的识别与处理能力却大为逊色,这说明尽管计算机的运算速度远远超过人脑,但是人脑的运算规则要优越得多。例如,“过马路”、“辨颜色”等,这些对于人类来讲轻而易举的事情,如果交给计算机来做,需要复杂的操作和运算,可见人脑要比计算机更聪明。因此,智慧的人类展开了揭示自身大脑之谜的研究。

现代神经生理学和神经解剖学的研究表明,人脑极其复杂,由约 10^{10} 个神经元交织在一起,构成一个网状结构。它能完成诸如思维、情绪等高级神经活动,被认为是最复杂、最完美、最有效的一种信息处理系统。人工神经网络(Artificial Neural Networks, 缩写 ANN)是对人脑若干基本特性通过数学方法进行抽象和模拟,是一种模仿人脑结构及其功能的非线性信息处理系统。

人工神经网络的产生与发展,经历了一个曲折艰难的过程。早在 1943 年,心理学家 Warren McCulloch 和数理逻辑学家 Walter Pitts 从信息处理的角度出发,采用数理逻辑的方法,对生物神经细胞的动作进行研究,提出了形式神经元模型,称为 MP 模型。通过 MP 模型提出了神经元的形式化数学描述和网络结构方法,证明了单个神经元能够执行逻辑功能,从而开创了人工神经网络研究的时代。1949 年,心理学家 Donald Hebb 明确提出了突触联系强度可变假设,认为突触联系强度是随前后神经元的活动而变化,从而说明了神经元连接强度是可修正的,为神经网络的产生奠定了基础。随后,感知器(Perceptron)、自适应线性神经元(Adaptive Linear Neuron, 简称 ADALINE)等网络与运行规则被相继提出,早期主要应用于模式识别、天气预报和自适应控制。然而,简单的感知器只能局限于处理线性问题,加之当时 Von Neumann 数字

计算机和人工智能正处于发展的黄金时期，暂时掩盖了发展新型计算机和人工智能新途径的必要性和迫切性，使人工神经网络的研究一度处于低潮。直至1982年，美国加州工学院物理学家 John Hopfield 提出了 Hopfield 神经网络模型，引入了“能量函数”的概念，给出了网络稳定性的判据，并成功地解决了著名的“旅行商问题”，成为人工神经网络走向成熟的里程碑。1986年，D. E. Rumelhart 和 J. L. McClelland 提出多层网络学习的误差反传播算法，为多层感知机找到了一个有效的学习算法，从而把人工神经网络的研究进一步推向深入。与此同时，基于 Von Neumann 原理的数字计算机在处理形象思维、语音图像识别以及联想记忆等模糊信息时屡受挫折，而神经网络能够处理模糊信息的特点使人们认识到人工神经网络的重要价值。随着人工神经网络理论研究的不断深入，神经网络的发展进入了一个新时期，应用研究也得到迅速发展，应用领域趋于广泛。进入20世纪90年代，人工神经网络作为新学科、新方法和新技术，在自然科学和社会科学各个领域得到了广泛的应用，取得了丰硕的成果。

遗传算法 (Genetic Algorithms) 是模拟生物界的遗传和进化过程而建立起来的一种搜索算法，体现着“生存竞争、优胜劣汰、适者生存”的竞争机制。早在20世纪50年代，自然界生物进化理论已被广泛接受。作为生物学家的 Fraser，试图通过计算的方法来模拟生物界“遗传与选择”的进化过程，这便是遗传算法的最早雏形。直到20世纪60年代，美国 Michigan 大学的 Holland 教授首次明确提出遗传算法的概念。Holland 创建的遗传算法，是基于二进制表达的概率搜索方法，在种群中，通过信息交换重新组合新染色体；根据评价条件，概率选择适应性好的染色体进入下一代；经过多代进化，种群最后稳定在适应性好的染色体上，即获得最优解。由于遗传算法的特点是种群搜索和种群中个体之间交换信息，搜索不依赖于梯度信息和问题本身，具有的鲁棒性和全局搜索能力，常常应用于优化和并行处理领域。

人工神经网络能够模仿人脑的结构及功能，具有对信息进行并行处理、分布式存贮以及自学习与推理的能力，表现出容错性、非线性、非局域性、非凸性等特点，适于对模糊信息或复杂的非线性关系进行识别与映射。遗传算法是模拟生物界遗传和进化过程而建立起来的一种搜索算法，具有高效的随机搜索与全局优化的特点，适于优化问题的求解。这些优良特性为采用神经网络理论与方法研究气象、水文以及水环境等问题奠定了坚实的理论基础。研究人工神经网络和遗传算法在水科学中的应用，丰富和发展了水科学的理论体系，增强了对非线性问题的识别与模拟能力，提高了处理问题的速度，为该领域研究注

入了新的活力，具有重要的理论价值和广泛的应用前景。

本书在系统阐述人工神经网络与遗传算法的基本概念、原理及运行过程的基础上，结合水科学领域的专业知识，紧密联系工程实际，全面、系统地介绍人工神经网络与遗传算法在气象、水文以及水环境等水科学领域的应用现状，深入研究了洪水预报、多泥沙洪水预报和水工建筑物的流激振动系统识别等实际预报和辨识模型及应用。根据防洪减灾实际问题，提出了人工神经网络峰值识别理论和遗传算法优化网络初始权重方法。

本书内容主要由三部分组成：

第一部分，介绍人工神经网络和遗传算法的基本知识。主要包括人工神经网络的产生与发展；人工神经网络的基本概念、特征与分类；常用网络的拓扑结构、工作原理、学习算法；遗传算法的基本概念和发展历程；遗传算法的运行过程及定理与假定；遗传算法的应用设计等。

第二部分，根据人工神经网络设计中经常遇到的问题，提出了人工神经网络峰值识别理论和遗传算法优化网络初始权重的实用方法，建立了将实际问题概化成神经网络可以识别的计算模型。

第三部分，阐述了人工神经网络在水科学中的应用成果。主要包括水系洪水预报、多泥沙洪水预报和流激振动系统的识别与预报等。

本书的创新点体现在：在理论方面，①采用人工神经网络技术解决水利领域的复杂、模糊、高度非线性问题；②提出了基于人工神经网络的峰值识别理论；③将遗传算法应用神经网络初始权重的优化，为避免网络陷入局部极小点提供了有效方法；④人工神经网络与遗传算法的有机结合，提高了智能算法的识别与映射能力。在实际应用方面，研制了基于人工神经网络与遗传算法理论的洪水智能预报系统和多泥沙洪水预报系统，进一步证实了智能理论在水科学应用中的可靠性和高效性。

本书是著者多年来将神经网络和遗传算法等现代智能研究成果应用于解决水科学领域实际问题的经验总结。希望本书的出版有利于进一步推动智能科学理论在水科学领域的广泛应用，有助于智能科学理论在实际应用中不断深化与发展。

本研究成果得益于国家自然科学基金项目的资助，得到了中国水利水电科学研究院、天津大学、珠江水利委员会和黄河水利委员会等单位以及有关专家、同仁的大力支持，在此表示诚挚的感谢。

著 者

2002年3月于北京

目 录

前 言

第一章 人工神经网络基本知识	1
第一节 人工神经网络简述	1
第二节 人工神经元模型	4
第三节 人工神经网络模型的拓扑结构	8
第四节 人工神经网络的运行过程	11
第五节 BP 网络的改进算法及优化设计	13
参考文献	26
第二章 遗传算法理论	29
第一节 概述	29
第二节 遗传算法的运行过程	31
第三节 遗传算法的定理与假设	49
第四节 遗传算法应用设计	58
参考文献	73
第三章 神经网络峰值识别理论和遗传算法优化权重方法	75
第一节 引言	75
第二节 数据信息的预处理	77
第三节 神经网络峰值识别理论	80
第四节 遗传算法优化 BP 网络初始化权重	83
参考文献	98
第四章 人工神经网络在降雨径流预报中的应用	99
第一节 降雨径流预报的基本方法	99
第二节 降雨径流的计算机仿真	100
第三节 小结	104
参考文献	104
第五章 人工神经网络在河道洪水预报中的应用	106
第一节 引言	106
第二节 洪水神经网络预报的基本原理	107
第三节 单一河道洪水神经网络预报	108
第四节 河系洪水神经网络预报	111

第五节 小结	113
参考文献	114
第六章 人工神经网络用于多泥沙洪水预报	117
第一节 多泥沙洪水预报的基本原理	117
第二节 多泥沙洪水预报的神经网络模型	118
第三节 水沙预报的计算机仿真	119
第四节 小结	124
参考文献	124
第七章 西江洪水神经网络预报	126
第一节 西江洪水概况	126
第二节 神经网络预报模型	127
第三节 模型训练及精度评定	131
第四节 预报模型的检验	140
第五节 小结	146
参考文献	147
第八章 流激振动系统的神经网络辨识和预报	148
第一节 人工神经网络辨识的内涵	148
第二节 黑箱振动系统的辨识	149
第三节 联合神经网络模型及算法	151
第四节 动态方程和流激振动系统的计算机仿真	154
第五节 人工神经网络方法的流激振动预报	159
参考文献	163
第九章 人工神经网络在相关领域中的应用	165
第一节 人工神经网络在气象科学领域的应用	165
第二节 水资源优化利用和预报	172
第三节 水环境识别与预报	173
参考文献	174

第一章 人工神经网络基本知识

第一节 人工神经网络简述

一、人工神经网络的基本涵义

现代神经生理学和神经解剖学的研究表明,人脑是极其复杂的,由约 10^{10} 个神经元交织在一起,构成一个网状结构。它能完成诸如智能、思维、情绪等高级神经活动,被认为是最复杂、最完美、最有效的一种信息处理系统。人工神经网络(Artificial Neural Networks, 缩写 ANN)是对人脑若干基本特性通过数学方法进行的抽象和模拟,是一种模仿人脑结构及其功能的非线性信息处理系统。

二、人工神经网络的产生与发展

人工神经网络的产生与发展经历了一个曲折艰难的过程。早在 1943 年,心理学家 Warren McCulloch 和数理逻辑学家 Walter Pitts 从信息处理的角度出发,采用数理模型的方法,对生物神经细胞的动作进行研究,提出了形式神经元模型,称为 MP 模型。通过 MP 模型提出了神经元的形式化数学描述和网络结构方法,证明了单个神经元能够执行逻辑功能,从而开创了人工神经网络研究的时代。1949 年,心理学家 Donald Hebb 在《行为科学》(The Organization of Behaviour)一书中,明确提出了突触联系强度可变假设,认为突触联系强度是随前后神经元的活动而变化,从而说明了神经元连接强度是可修正的,为神经网络的产生奠定了基础。1957 年 Frank Rosenblatt 首次引入感知器(Perceptron)的概念,它由阈值型神经元组成,试图模拟动物和人脑的感知和学习能力。感知器通过训练可用作某些模式的分类器。尽管感知器模型十分简单,但它已经具备了神经网络的某些基本特征,如学习功能、分布式存贮和并行处理功能、一定的容错性等等。1962 年,美国斯坦福大学教授 Bernhard Widrow 提出了自适应线性神经元(Adaptive Linear Neuron, 简称 ADALINE),并与 Marcian Hoff 一起提出了一个新的学习规则,称为 Widrow-Hoff 学习规则。此规则可减少训练过程中神经网络输出的误差平方和。自适应线性神经元及其扩展而来的多自适应线性神经元(Many ADALINE, 简称 MADALINE),早期主要应用于模式识别、天气预报和自适应控制。人工神经网络第一个成功应用的实例是消除通讯中的回波。

但是,在 1969 年 Marvin Minsky 和 Seymour Papert 出版了《感知器》(Perceptron)一书,该书指出:简单的感知器只能局限于处理线性问题,对于像 XOR(异或)这样简单的非线性问题都无能为力。他们提出了构造含有隐层的多层感知器的方案,但对隐层神经元的学习规则并不清楚,并对此持悲观态度。加之当时 Von Neumann 数字计算机和人工智能正处于发展的黄金时期,暂时掩盖了发展新型计算机和人工智能新途径的必要性和迫切性,使人工神经网络的研究一度处于低潮。

直至 1982 年,美国加州工学院物理学家 John Hopfield 提出了 Hopfield 神经网络模型,

引入了“能量函数”的概念，给出了网络稳定性的判据，并成功地解决了著名的“旅行商问题”，成为人工神经网络走向成熟的里程碑。1984年，他又提出了连续时间 Hopfield 神经网络模型，为神经计算机的研究做了开拓性的工作，开创了神经网络用于联想记忆和优化计算的新途径，有力地推动了神经网络的研究。

1986年，D.E. Rumelhart 和 J.L. McClelland 提出多层网络学习的误差反传播算法，实现了 Minsky 引入隐层的设想。这为多层感知机找到了一个有效的学习算法，从而把人工神经网络的研究进一步推向深入。与此同时，基于 Von Neumann 原理的数字计算机在处理形象思维、语音图像识别以及联想记忆等模糊信息时屡受挫折，而神经网络能够处理模糊信息的特点使人们认识到人工神经网络的重要价值。

随着人工神经网络理论研究的不断成熟，神经网络的发展进入了一个新时期，应用研究也得到迅速发展，应用领域趋于广泛。进入 20 世纪 90 年代，人工神经网络作为新学科、新方法和新技术，在自然科学和社会科学各个领域得到了广泛的应用，取得了丰硕的成果。神经网络研究也随之得到飞速发展，各种网络结构和算法系统应运而生，逐渐构成了较为完善的人工神经网络理论体系。同时，人工神经网络的学术交流日益频繁，形成了良好的学术氛围。

1987年6月电气和电子工程师协会（IEEE）在美国加州举行了第一届人工神经网络（ANN）国际会议，并成立了国际神经网络学会，此后，每年召开两次国际联合神经网络大会（IJCNN）。我国于1990年12月在京召开了第一次中国神经网络学术大会，规定每年召开一次，并于1991年成立了中国神经网络学会。

人工神经网络作为一种数学算法体系，已经解决了许多实际问题，它的生命力也恰恰在于其广泛的实用价值。

三、人工神经网络的特点

与传统的 Von Neumann 计算机相比较而言，人工神经网络具有以下几个突出的优点：

1. 大规模的并行计算与分布式存贮能力

传统计算机的计算和存贮是互相独立的，而在 ANN 中，无论是单个神经元还是整个神经网络都兼有信息的处理和存贮的双重功能，这两种功能自然融合在同一网络中。ANN 计算过程的并行性决定了其对信息的高速处理能力。

2. 非线性映射能力

ANN 各神经元的映射特征是非线性的，有些网络的单元间采用复杂的非线性连接。因此，ANN 是一个大规模的非线性动力系统，具有很强的非线性处理能力。

3. 较强的鲁棒性和容错性

由于信息的分布存储和集体协作计算，每个信息处理单元既包含对集体的贡献又无法决定网络的整体状态，因此，局部神经网络的故障并不影响整体神经网络输出的正确性。

4. 自适应、自组织、自学习的能力

神经网络最突出的特点就是具有自适应、自组织、自学习的能力，它可以处理各种变化的信息，而且在处理信息的同时，非线性动力系统本身也在不断变化，即可以通过对信息的有监督和无监督学习，实现对任意复杂函数的映射，从而适应环境的变化。

5. 非局域性

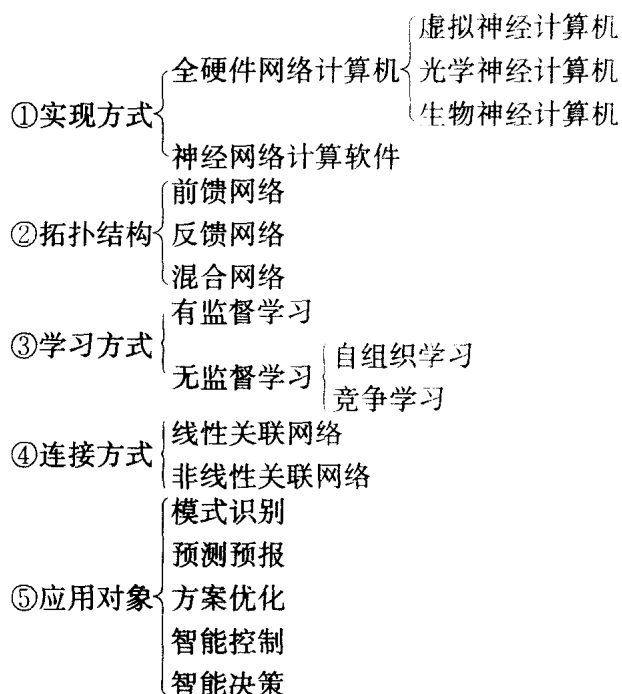
一个神经网络通常由多个神经元广泛连接而成。一个系统的整体行为不仅取决于单个神经元的特征，而且可能由神经元之间的相互作用、相互连接所决定，通过神经元之间的连接模拟大脑的非局域性，联想记忆是非局域性的典型例子。

6. 非凸性

一个系统的演化方向，在一定条件下将取决于某个特定的状态函数，例如能量函数，它的极值对应系统某个比较稳定的状态。非凸性是指某系统的能量函数有多个极值，故系统具有多个较稳定的平衡状态，这将导致系统演化结果的多样性。

四、人工神经网络的分类

人工神经网络分类标准不同，方法多样，归纳起来可分为以下几种形式：



还有根据神经网络算法的不同进行分类的，如 BP 网络、Hopfield 网络、Boltzmann 机、Gauss 机、Cauchy 机等。

五、人工神经网络的广泛应用

随着计算机技术的普及与发展，人工神经网络的研究与应用也取得了惊人的进展，涉及自然科学、社会科学、应用科学及综合性交叉科学等各个方面，取得了令人瞩目的成果，主要的应用有以下几个方面。

1. 模式识别

在模式识别方面的主要应用有：非线性系统识别、动态系统识别、图形和文字识别、卫星遥感图像识别、声音识别、模糊系统识别、化学过程的仿真与识别、医疗疾病的诊断等。

2. 预测和预报

在预测和预报方面的主要应用有：经济发展预测、银行汇率预测、股市行情预测、市

场需求预测、天气预报、降雨径流预报、洪水预报、电力负荷预测、交通状况预测等。

3. 优化问题

电力输送的优化、大型混合问题的解决、能源的优化利用、水资源优化配置、水库优化调度、大型拱坝体型优化等。

4. 神经控制

基于神经网络的智能控制有：①工业过程控制，包括蒸汽机的模糊控制、退火炉燃烧过程控制、冶炼炉的模糊控制、造纸过程控制、仿人智能温度控制器、电力系统控制；②运载工具控制，包括飞机机翼的倾斜和力矩控制、列车自动驾驶系统的模糊智能控制、汽车喷油系统的神经网络控制；③机器人的智能控制；④冰箱温度智能控制系统；⑤机械振动系统的智能减振控制；⑥大型结构的智能减振控制等。

5. 智能决策和专家系统

美国 BehavHeuristics 公司应用自适应神经网络研制出一整套航空管理与决策专家系统，实现航空收入的管理、航空市场的分析和预测、人力物力的优化分配与调度；智能数据库系统；专家系统中智能学习和判断分析系统等。

第二节 人工神经元模型

人工神经元是人工神经网络的基本单元，正如人脑的脑细胞。它模拟了脑细胞对信息刺激的反映。我们先从生物神经元的结构与基本特性入手，逐步介绍人工神经元模型的建立。

一、生物神经元的结构及其基本特征

生物解剖学的研究表明，人脑由 10^{10} 数量级的称为神经元的脑细胞组成，神经元彼此连接，构成了空间上的网络结构。尽管不存在两个完全相同的神经元，但这些神经元的结构与功能是大体相同的。图 1-1 为一个典型生物神经元的结构示意图。

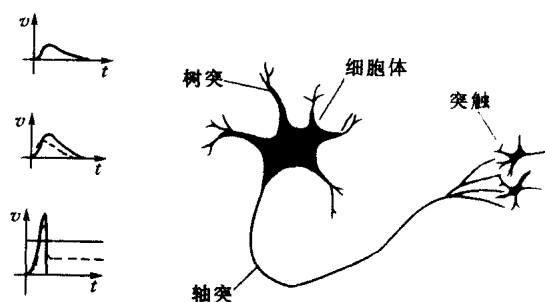


图 1-1 生物神经元结构示意图

生物神经元由细胞体、树突和轴突构成，每个细胞只有一个细胞核、一根轴突，数以千万计的树突及其突触。

轴突粗细均匀，表面光滑，其功能为传输细胞体发出的神经信息，是该神经元的输出通道。

树突粗而短，数量非常多，其作用为收集其他神经元传出的信息，是神经元的输入通道。

突触是两个神经元连接之处，各神经元间通过突触完成信息的传递，而且信息是单向的。如果通过某个突触传来的信号起到使该神经元兴奋的作用，则称其为兴奋型突触；反之，称抑制型突触。另外，突触的连接强度代表了两个神经元间信号传递时耦合的紧密程度，不同的紧密程度相当于作用信号不同的“权重”，而且，正权重表示兴奋作用，负权重表示抑制作用。所有输入信号加权后的

综合效果作为神经元的输入。若输入值大于或等于该神经元的阈值，该神经元便被激活；否则，未被激活。

神经信号沿着轴突传出，激励突触连接，接着产生下一个激励，沿着树突传入另一个神经元。

综上所述，生物神经元具有如下基本特性：

- (1) 神经元具有多输入、单输出的特性。
- (2) 神经元具有非线性输入、输出的特性。
- (3) 各神经元间传递信号的强度是可变的，输入的信号有兴奋作用与抑制作用之分。
- (4) 神经元的输出响应取决于所有输入信号的加权效果，当等效的输入超过某一阈值时，该神经元被激活；否则，处于抑制状态。

二、生物神经元的模拟

概括起来，生物神经元的处理信息包括两个阶段，第一阶段是神经元接收信息流的加权过程，称作聚合过程；第二阶段是对聚合后信息流的线性、半线性、非线性函数的处理过程，称作活化过程。不同的信息处理函数反映了神经元处理复杂信息能力的差异。将输入信息和响应输出用一个传输方程表示为：

$$Y = F\left(\sum_{k=1}^n X_k W_k\right) \quad (1-1)$$

式中 Y ——神经元的输出；

F ——神经元对输入信息的响应特性；

X_k ——第 k 个突触所对应的输入信息；

W_k ——第 k 个突触的权值。

图 1-2 模拟了单个生物神经元处理信息的过程，这一过程还可以简化成图 1-3 的形式。

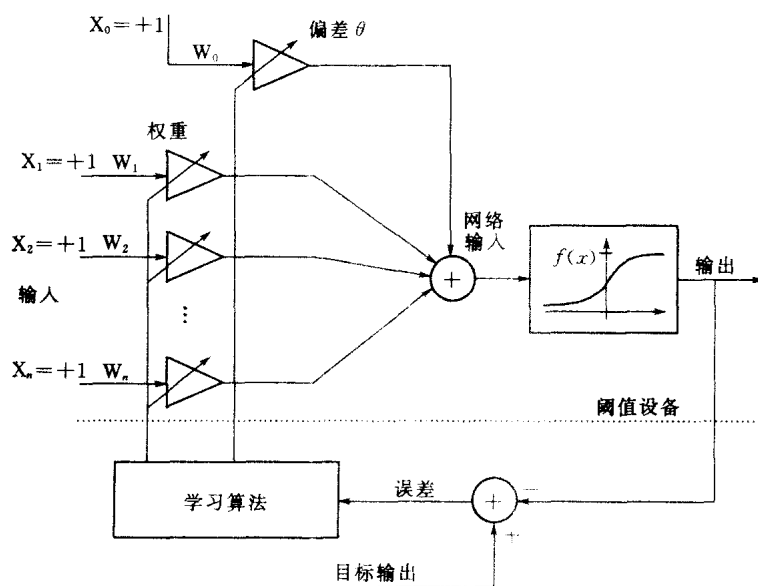


图 1-2 神经元的处理信息过程模拟

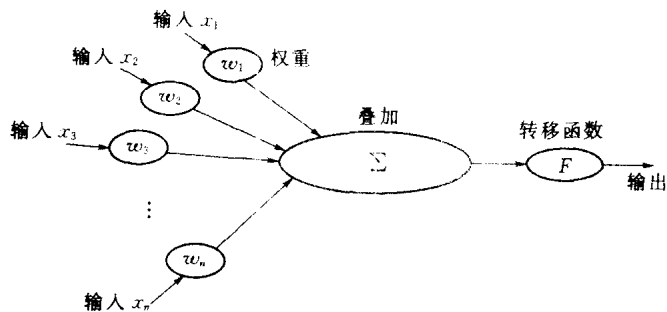


图 1-3 神经元的信总处理过程简化示意图

单个神经元信息识别的能力是有限的，只能反映极其简单的信息过程，只有将多个神经元组合起来，构成一个神经网络体系，才能处理和识别复杂的信息流变化过程和相关机制。

三、人工神经元的数学模型

将图 1-3 神经元的信总处理过程采用数学方式进行描述，便得到了如图 1-4 所示的人工神经元的数学模型。

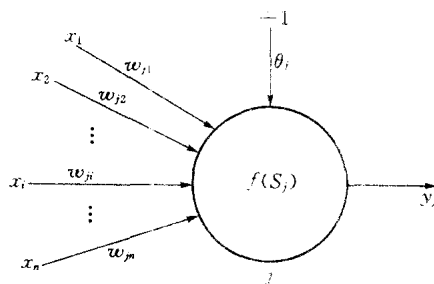


图 1-4 人工神经元的数学模型

这里的 $x_1, x_2, \dots, x_i, \dots, x_n$ 分别表示来自其他神经元突触的输入，相应的 $w_{j1}, w_{j2}, \dots, w_{ji}, \dots, w_{jn}$ 表示突触 1, 2, $\dots, i, \dots, n$ 与第 j 个神经元的突触连接强度，即连接权重。

人工神经元的数学模型描述了一个典型生物神经元，是对信息进行聚合处理和活化处理的完整生物反应过程。需要指出的是，数学模型只是对生物神经元的简单模拟，忽略了生物神经元响应的时间延迟，而是一有输入便立即产生输出；此外，也没有考虑生物神经元的频率调制功能。

为了方便分析，我们把整个过程分解为三个数学计算步骤：

加权——对每个输入信号进行程度不等的加权计算；

求和——进行全部输入信号的组合效果的求和计算；

映射——通过转移函数 $f(\cdot)$ 计算输出结果。

下面以数学公式描述生物神经元响应过程。

列向量 $\mathbf{X}$ 表示输入向量为：

$$\mathbf{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = [x_1 \quad \cdots \quad x_i \quad \cdots \quad x_n]^T \quad (1-2)$$

行向量 $\mathbf{W}_j$ 表示神经元 j 的连接权重向量为:

$$\mathbf{W}_j = [w_{j1} \quad w_{j2} \quad \cdots \quad w_{ji} \quad \cdots \quad w_{jn}] \quad (1-3)$$

神经元 j 的净输入 S_j 为:

$$S_j = \sum_{i=1}^n w_{ji}x_i + \theta_j = \mathbf{W}_j\mathbf{X} + \theta_j \quad (1-4)$$

式中 θ_j ——阈值 (Threshold)。

净输入经过转移函数的作用后, 得到神经元的输出 y_j 为:

$$y_j = f(S_j) = f\left(\sum_{i=1}^n w_{ji}x_i + \theta_j\right) \quad (1-5)$$

若 θ_j 采用 $x_0=1$ 的处理方法, 则公式 (1-5) 表示为:

$$y_j = f(S_j) = f\left(\sum_{i=0}^n w_{ji}x_i\right) = f(\mathbf{W}_j\mathbf{X}) \quad (1-6)$$

至此, 用公式的形式描述了一个神经元由接收信号到作出反应的信息处理过程。

转移函数 $f(\cdot)$ 又称激活函数 (Activation Function), 其作用是模拟生物神经元所具有的非线性转移特性。线性函数和对数形式的 Sigmoid 函数是最为常用的两种转移函数。

1. 线性函数

$$y = f(x) = x \quad (1-7)$$

2. Sigmoid 函数 (S 型曲线)

其特点是: ①有上、下界; ②单调增函数; ③连续且光滑, 即可微分。常用的 Sigmoid 函数是对数函数为:

$$y = f(x) = \frac{1}{1 + e^{-tx}} \quad (1-8)$$

从图 1-5 中可以发现, 对数型的 Sigmoid 函数上下限分别为 0 和 1, 并且系数 t 决定

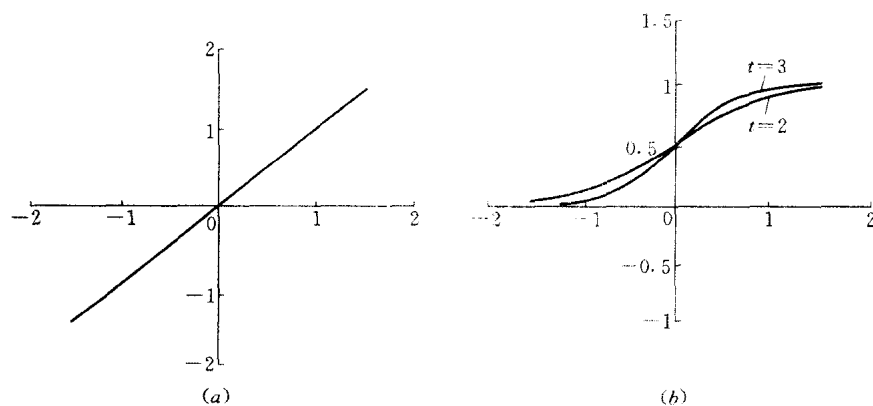


图 1-5 常用的转移函数
(a) 线性函数; (b) Sigmoid 函数

了函数的压缩程度。系数 t 越大，曲线越陡；反之，曲线越缓。

第三节 人工神经网络模型的拓扑结构

前面模拟了单个生物神经元对信息的处理过程，建立了人工神经元模型。单个神经元对信息的处理能力是非常有限的，只有将多个神经元相互连接起来，构成一个神经网络体系，才能够对复杂的信息进行识别处理。这里，将介绍神经元的连接方式，即人工神经网络的拓扑结构。

根据神经网络的拓扑结构和信息流在其中的传递方式，人工神经网络可以大致分为前馈网络、反馈网络和混合网络三种形式。

一、前馈网络

前馈网络 (Feedforward Networks) 的信息流由输入层，逐级向下层传递，没有反馈信息流，经网络处理后由输出层输出。单层前馈网络是最简单的前馈式网络，如图 1-6 所示。对于多层前馈网络，输入、输出神经元与外界发生联系，直接感受外部环境的刺激；而中间层与外界无直接联系，所以称为隐层。图 1-7 为多层前馈网络的拓扑结构示意图。

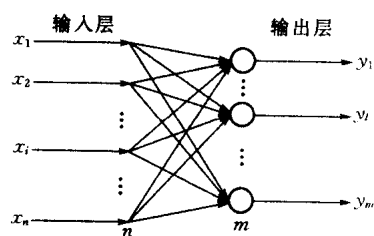


图 1-6 单层前馈网络

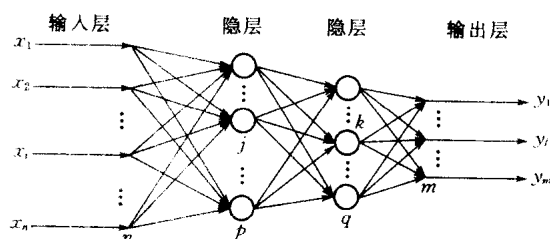


图 1-7 多层前馈网络拓扑结构

这里，以单层前馈网络为例，描述人工神经网络对信息流的处理过程。

如果，网络的输入列向量 $\mathbf{X}$ 为：

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = [x_1 \quad x_2 \quad \cdots \quad x_i \quad \cdots \quad x_n]^T \quad (1-9)$$

神经元节点的连接权重表示为权重矩阵 $\mathbf{W}$ ，权重矩阵 $\mathbf{W}$ 的行数等于输出神经元节点数 m ，列数等于输入神经元节点数 n 。 $\mathbf{W}$ 的 j 行是神经元节点 j 的权重向量 $\bar{\mathbf{W}}_j$ (行向量)。

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1i} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2i} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{j1} & w_{j2} & \cdots & w_{ji} & \cdots & w_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mi} & \cdots & w_{mn} \end{bmatrix}_{m \times n} = \begin{bmatrix} \bar{w}_1 \\ \bar{w}_2 \\ \vdots \\ \bar{w}_j \\ \vdots \\ \bar{w}_m \end{bmatrix} \quad (1-10)$$

阈值 θ 为:

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_j \\ \vdots \\ \theta_m \end{bmatrix} = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_j \quad \cdots \quad \theta_m]^T \quad (1-11)$$

神经元节点的净输入 S 为:

$$S = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_j \\ \vdots \\ s_m \end{bmatrix} = [s_1 \quad s_2 \quad \cdots \quad s_j \quad \cdots \quad s_m]^T \quad (1-12)$$

则有:

$$WX + \theta = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1i} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2i} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{j1} & w_{j2} & \cdots & w_{ji} & \cdots & w_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mi} & \cdots & w_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_j \\ \vdots \\ \theta_m \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_j \\ \vdots \\ s_m \end{bmatrix} = S \quad (1-13)$$

若引入矩阵符号 Γ 为:

$$\Gamma = \begin{bmatrix} f(\cdot) & 0 & \cdots & 0 \\ 0 & f(\cdot) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & f(\cdot) \end{bmatrix}_{m \times m} \quad (1-14)$$

来表示转移函数对神经元节点净输入的转移作用, 则: