

多层前向

编 王 魏 胡 赵
树 德 浩 林
著 谦 华 云 明

人工神经网络

- 基本理论
- 学习算法
- 各种类型
- 应用举例

QXRG SJWL

黄河水利出版社

多层前向人工神经网络

赵林明 胡浩云 魏德华 王树谦 编著

黄河水利出版社

图书在版编目 (CIP) 数据

多层前向人工神经网络/赵林明等编著. - 郑州: 黄河水利出版社, 1999.7
ISBN 7-80621-215-9

I. 多… II. 赵… III. 人工神经网络 IV. TP18

中国版本图书馆 CIP 数据核字 (1999) 第 14104 号

责任编辑: 荆东亮

封面设计: 朱 鹏

责任校对: 赵宏伟

责任印制: 常红昕

出版发行: 黄河水利出版社

地址: 河南省郑州市顺河路黄委会综合楼 12 层 邮编: 450003

E-mail: yrep@public2. zz. ha. cn

印 刷: 黄委会设计院印刷厂

开	本: 850mm × 1168mm	1/32	印	张: 4.625
版	别: 1999 年 7 月	第 1 版	印	数: 1 - 1 000
印	次: 1999 年 7 月	郑州第 1 次印刷	字	数: 115 千字

定价: 10.80 元

前 言

人工神经网络 (Artificial Neural Network: 简记 ANN) 是在研究生物神经系统的启发下发展起来的一种信息处理方法, 具有学习联想能力等特点。目前全球范围内, 在信息处理、智能控制、模式识别等许多工程技术和社会经济领域得到了越来越多的应用。最近几年来, 我国水电工程技术领域的科学技术人员也将人工神经网络方法应用于水力发电过程辨识和控制、河川径流预测、河流水质分类、水资源规划、混凝土性能预估、电力系统分析、拱坝优化设计、预应力混凝土桩基等结构损伤诊断、砂土液化预测、岩体可爆性分级及爆破效应预测、岩土类型识别、地下工程围岩分类及支护决策、大坝等工程结构安全监测、工程造价分析等许多实际问题中。

目前已出版了一些有关人工神经网络的书籍, 但在这些书籍中基本上是全面介绍了人工神经网络的有关内容。而在水利水电工程领域中, 应用神经网络技术解决具体问题时, 基本上采用的是多层前向人工神经网络, 网络的训练过程是采用有导师的学习方法。针对这一情况, 根据编著者的研究应用与教学实践, 并在综合近年来发表的大量有关文献资料的基础上, 以多层前向人工神经网络为主要内容, 从应用角度出发, 编写了本书, 以推动人工神经网络技术的应用和普及。

本书由赵林明等编著。全书共分五章, 分别介绍了人工神经网络的一般概念, 多层前向人工神经网络的基本原理, 多层前向人工神经网络的典型学习方法, 各种类型的多层前向人工神经网络, 以及应用多层前向人工神经网络解决水电工程中一些问题的实例, 并附有反向传播算法的计算程序。其中第三章第一、二节

和第五章第三、五节及附录由胡浩云编写；第三章第五、六、七节和第四章第三、七、八节及第五章第二、七节由魏德华编写；第三章第三节和第四章第一、二、四、五节及第五章第四、六、八节由王树谦编写；赵林明编写了其余部分，并对全书进行了统稿。

为了便于读者掌握、理解和应用本书中的内容，本书最后列出了一些主要的参考文献，在此对各文献的作者表示诚挚的谢意。

书中给出的内容，可供水利水电工程及其他领域的科技人员学习使用，也可供水利水电工程及其他学科的研究生和本科高年级学生作为教材使用。由于编著者水平有限，书中不妥之处在所难免，恳望读者批评指正。

编著者

1998 年 12 月

目 录

第一章 概述	(1)
1.1 大脑神经网络	(1)
1.2 人工神经网络的组成	(4)
1.3 人工神经网络的发展简史	(7)
第二章 多层前向人工神经网络的基本理论	(9)
2.1 多层前向人工神经网络的基本组成	(9)
2.2 多层前向人工神经网络的非线性逼近能力	(12)
2.3 多层前向人工神经网络的学习过程	(14)
第三章 多层前向人工神经网络的学习算法	(16)
3.1 学习样本的收集和整理	(16)
3.2 反向传播学习算法	(23)
3.3 全局优化学习算法	(38)
3.4 模拟进化优化学习算法	(46)
3.5 模拟退火学习算法	(57)
3.6 交替迭代学习算法	(63)
3.7 广义递推最小二乘学习算法	(67)
第四章 多层前向人工神经网络的各种类型	(69)
4.1 双并联前向人工神经网络	(69)
4.2 深层前向人工神经网络	(73)
4.3 自适应前向人工神经网络	(76)
4.4 模糊前向人工神经网络	(82)
4.5 非全连接的前向人工神经网络	(85)
4.6 高阶前向人工神经网络和含有虚拟输入的	

前向人工神经网络	(88)
4.7 径向基函数前向人工神经网络	(91)
4.8 埃曼前向人工神经网络	(94)
第五章 多层前向人工神经网络在水利水电工程中的应用	(97)
5.1 在建立水轮机输入输出关系数学模型中的应用	(97)
5.2 在水力发电过程控制中的应用	(102)
5.3 在建立河流水质评价模型中的应用	(105)
5.4 在水库优化调度中的应用	(108)
5.5 在混凝土性能分析和预测中的应用	(110)
5.6 在水文分类预报中的应用	(111)
5.7 在电力系统分析中的应用	(114)
5.8 在湖泊富营养化程度评价中的应用	(119)
附录 反向传播算法计算程序	(122)

第一章 概 述

目前得到广泛应用的人工神经网络 (Artificial Neural Network: 简记 ANN), 又称连接机制模型 (Connectionism Model) 或并行分布处理模型 (Parallal Distributed Model), 是由大量简单的元件连接而成的, 用以模拟人脑行为的复杂网络系统。人工神经网络反映了人类大脑功能的许多基本特征, 但并不是人脑神经网络系统的真实写照, 而只是对其作某种简化、抽象和模拟。本章简要地介绍一下人工神经网络的基本概念和发展简史, 需要进一步了解人工神经网络的详细内容的读者, 可以参阅书后参考文献 [1~6]。

1.1 大脑神经网络

为了更好地理解人工神经网络的组成和特性, 下面简要介绍一下人脑组织模型。

1.1.1 神经元

大脑的神经系统是称为脑细胞的 10^{11} 到 10^{12} 个神经元组成, 每个神经元可直接与 1 000 到 20 000 个其他神经元相互作用。大脑强大的功能来自大量的多层次的神元元的相互连接。

最简单的生物神经元模型如图 1.1 所示, 一神经元的中心有一实体团, 称为细胞核, 它能对接收到的信息进行处理。细胞核周围的纤维分为两类, 一类为接收信息的称为树突, 另一类是发

出信息的称为轴突。

树突可以看作细胞体的延伸部，它自细胞体发出后逐渐变细，其全长各点都可接受其他神经元的轴突末梢，形成突触。突触是一个神经元与另一个神经元间发生功能联系的界面，并不是两神经元的细胞质的连通。突触分为化学性突触和电突触。

神经元的大小、形状以及突触的数目和分支方式差别很大，通常分为单级、双级和多级神经元等几种类型。

树突从突触接收到信号后，把信号引导到神经元的细胞核，信号在那里积累，激起神经元兴奋或抑制，从而决定了神经元的状态，两神经元结合部的突触，决定了神经元之间相互作用的强弱。

人工神经网络中的人工神经元（有时也称为计算单元或计算节点），模拟了大脑神经元的这些简单特性。

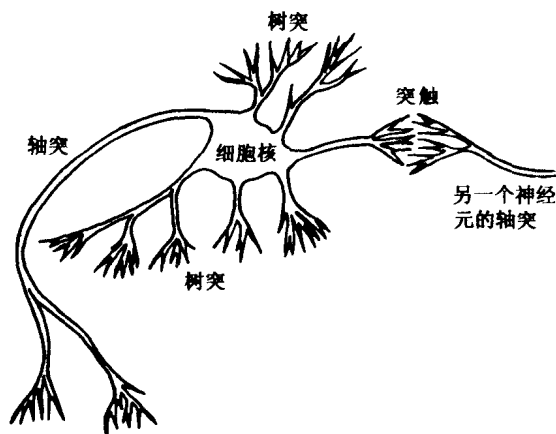


图 1.1 神经元的构造

1.1.2 大脑神经网络系统

以上简要介绍了神经元的结构及其机能。大脑中的神经元通

过突触互相连接，形成了复杂的神经网络系统，实现人类的各种智能行为。

作为一个智能的信息处理系统，人脑具有以下几方面的固有特征。

第一个特征是并行分布处理的工作模式。单个神经元的处理速度很慢，每次约 1ms，但人脑对某一复杂过程的反应却是很快，一般只需几百微秒。例如，要判断人眼看到的两个图形是否一样，实际需要 400ms。在这个处理过程中，与脑神经系统的一些主要功能，如视觉、记忆、推理等有关。如果采用串行工作模式，就没有如此快的反应能力。因此，只能把人脑看做一个超高密度的并行处理系统。

脑信息处理的第二个特征是神经系统的可塑性和自组织性。人工神经网络的学习机制就是基于这种可塑现象，并通过修正突触的结合强度来实现的。

脑信息处理的第三个特征是神经系统的系统性。大脑中的各个部件可以看成是一个大系统中的许多子系统，各个子系统之间具有很强的相互关系，一些子系统可以调节另一些子系统的行为。例如视觉系统和运动系统就存在很强的系统联系，可相互协调各种信息处理功能。

最后一个特征是系统的恰当退化。脑神经网络的信息存储是分布存储于整个网络中，并体现在神经网络之间突触的结合强度上。当少量神经元受到损伤（或正常死亡）时，网络的总体功能继续有效。

人工神经网络以脑神经系统的组成原理为构造基础，模拟了神经信息处理系统的上述四个基本特征，具有广泛的应用价值。

1.2 人工神经网络的组成

1.2.1 人工神经元

根据生物神经元的结构、作用机制，并作了进一步的简化，构成了神经元模型，即人工神经元，简称神经元，其结构如图 1.2。

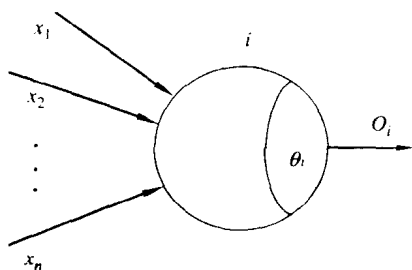


图 1.2 人工神经元模型

每一个神经元（如神经元 i ）接受其他神经元（如神经元 j ）的信息传递，总输入为

$$I_i = \sum_{j=1}^n w_{ij} x_j + \theta_i \quad (1.1)$$

式中： w_{ij} 表示神经元 i 和神经元 j 的结合强度，称为连接权； x_j 为神经元 j 的输出； θ_i 表示神经元 i 的阈值。

神经元 i 的输出为

$$O_i = f(I_i) \quad (1.2)$$

这里函数 $f()$ 称为激发函数，根据 $f()$ 的不同，将神经元模型分为不同的类型，如离散型、连续型、微分/差分型和概率型。

在离散输出模型中，激发函数选用阶跃函数，如

$$O_i = \begin{cases} 1 & I_i > 0 \\ 0 & I_i \leq 0 \end{cases} \quad (1.3)$$

应用最为广泛的是连续型，在连续输出模型中采用的激发函数为连续函数，如多项式函数、三角函数、样条函数等。目前广为采用的是 Sigmoid 函数，简称 S 型函数，即取神经元的输出为

$$O_i = \frac{1}{1 + e^{-I_i}} \quad (1.4)$$

Sigmoid 函数的图像如图 1.3。

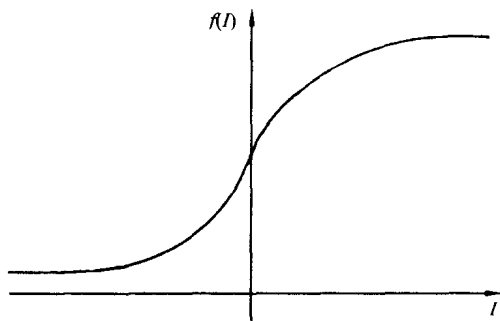


图 1.3 Sigmoid 函数的图像

S 型函数具有以下一些良好的特性：

(1) 当 I 较小时，也有一定的 O 值相对应，即输入到神经元的信号比较弱时，神经元也有输出，这样不丢失较小的信息反映；

(2) 当 I 较大时，输出趋于常数，不会出现“溢出”现象；

(3) 具有良好的微分特性，即有

$$\left(\frac{dO}{dI} \right)_i = O_i (1 - O_i) \quad (1.5)$$

由于有上述原因，S 型函数作为神经元的激发函数，得到了广泛的应用。为了进一步提高神经网络的实用性，近年有的文献介绍，将 S 型函数改进为如下形式

$$O_i = \frac{1}{a + be^{-d_i}} \quad (1.6)$$

式中： a 、 b 和 c 为可调参数。通过改变系数 a 、 b 和 c 的值，改变了图 1.3 中 S 型函数图像的几何形状，也就是改变了相应神经元的输入输出特性。

1.2.2 人工神经网络

上面介绍的是单个神经元模型，将许多神经元按照一定的方式组合成一个网络，就构成了人工神经网络。除网络中单个神经元的特性外，网络的拓扑结构也对网络的特性有重要的影响。根据网络中神经元的连接方式，将网络分成以下两种基本型式。

1.2.2.1 前向网络

目前在水电工程领域，广泛采用的是前向网络。前向网络中，各神经元接受前一级输入，单输出到下一级，无反馈。其基本结构如图 1.4，本书主要给出这种网络的有关内容及应用。

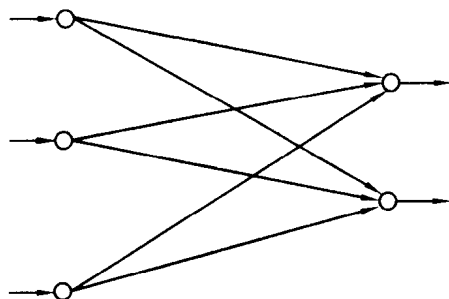


图 1.4 前向网络

1.2.2.2 反馈网络

所有神经元都是计算单元，同时可接受输入，并向外界输

出，其结构如图 1.5。图中，双向箭头表示神经元之间的信息是双向传递的。

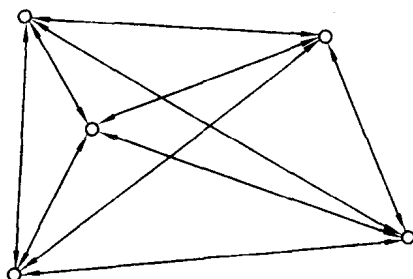


图 1.5 反馈网络

由以上两种基本网络，可以混合构成多种形式的网络，这里就不仔细讨论了。

人工神经网络在许多领域得到了应用，目前主要由软件来实现。部分情况下，采用硬件实现，即用一些电子的模拟器件（如运算放大器等）实现网络的组成，以提高运算速度，最终的努力目标是神经网络计算机。

1.3 人工神经网络的发展简史

1943 年，由心理学家 W. McCulloch 和数理逻辑学家 W. Pitts 首次提出了人工神经网络模型。在此之后，各国的专家学者相继对这一领域开展了各种研究，并取得许多成果。其中比较著名的研究成果是 Frank Rosenblatt 发展了 W. McCulloch 和 W. Pitts 提出的模型，并给出了感知器（Perceptron）模型（1957 年）以及两层感知器的收敛定理（1962 年），提出了引入隐层处理元件的三层感知器这一重要的研究方向。1969 年，美国的 Papert 教授和 Minsky 教授出版了《Perceptrons》一书，对感知器的功能及其局限性

从数学上作了深入的研究，提出了双层感知器的许多局限性。由于 Minsky 教授的个人威望及证明的严格性，使相当多的人认为人工神经网络的研究前途渺茫。此后，有关研究处于低潮，用于神经网络研究的经费大减。

1982 年，神经网络的研究有了转折，物理学家 John Hopfield 提出了一个用于联想记忆和优化计算的 HNN 模型，克服了 Minsky 提出的局限性，使人们对人工神经网络有了新的认识，一大批科学家又在这一领域开展了新的研究，推动了人工神经网络的发展。1987 年，国际电气和电工工程师协会（IEEE）举行首次神经网络国际会议，1 800 多人出席。同年，国际神经网络学会成立。现在，全球性的人工神经网络研究热潮方兴未艾，神经网络理论的应用已经渗透到各个领域，在计算机科学、脑神经科学、认知科学、心理学、微电子学、控制论、信息技术、数学和力学等方面取得了令人鼓舞的研究成果。美国工程索引（EI）每年摘录有关神经网络方面的文章都在 6 000 篇以上。

在近十几年来，我国的科技工作者也开展了大量的有关人工神经网络的研究和应用工作，其中在水利水电工程领域，近年来也取得了多项研究成果，本书最后一章中将给出一些实例。

第二章 多层前向人工神经网络的基本理论

正如在第一章中所述,多层前向人工神经网络是人工神经网络的一种主要类型,在这一章中,主要论述一般三层前向人工神经网络的有关理论。更多层次的前向人工神经网络和各种类型的多层前向人工神经网络将在第四章中详细论述。

2.1 多层前向人工神经网络的基本组成

2.1.1 三层前向人工神经网络的结构

目前得到广泛应用的多层前向人工神经网络是三层前向人工神经网络,三层前向人工神经网络由输入层、隐层和输出层组成,如图 2.1。

在三层前向人工神经网络中,输入层中有 n 个神经元,隐层中有 m 个神经元,输出层中有 l 个神经元。在一般情况下,整个网络的输入分别为 $x_1, x_2, \dots, x_n$, 且对应输入到网络输入层的第 1, 2, $\dots, n$ 个神经元,输入层中各神经元的激发函数一般都选用比例系数为 1 的线性函数,因此,输入层中神经元的输出信息分别是 $x_1, x_2, \dots, x_n$, 这些输出信息传递到隐层各神经元。隐层神经元将输入层输出的信息按照一定的方式汇总,作为各神经元的输入信息,然后通过激发函数的作用,产生隐层各神经元的输出信息。隐层神经元的激发函数一般选用非线性函数,隐层神经元的输出又传递到输出层,与隐层神经元类

似，输出层神经元也输出相应的信息，并作为整个网络的输出信息 $y_1, y_2, \dots, y_l$ 。

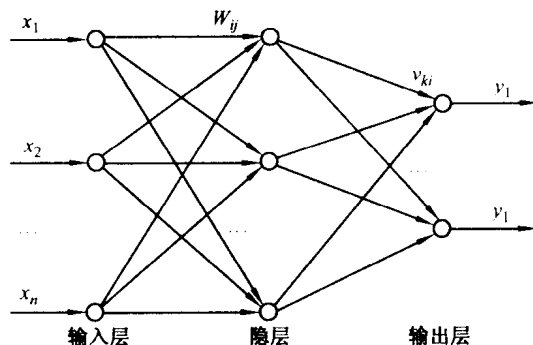


图 2.1 三层前向人工神经网络

从上述内容可知，输入到网络的信息是按照输入层、隐层、输出层的顺序，一直传递到输出层，最后得到整个网络的输出，因此，整个网络叫做前向网络。

2.1.2 前向人工神经网络的计算过程

上面给出了前向人工神经网络中信息的传递过程，下面给出具体的计算过程。

如上面内容中所述，网络中输入层的输出分别是 $x_1, x_2, \dots, x_n$ 。根据人工神经元的基本原理，隐层各神经元的输入分别是

$$I_i = \sum_{j=1}^n w_{ij} x_j + \theta_i \quad (i = 1, 2, \dots, m) \quad (2.1)$$

式中： w_{ij} 为隐层神经元 i 与输入层神经元 j 的连接权； θ_i 为隐层神经元的阈值，选择一定的函数作为隐层神经元的激发函数 $f()$ ，则隐层神经元的输出为

$$O_i = f(I_i) \quad (i = 1, 2, \dots, m) \quad (2.2)$$