

XIAN JIAOTONG UNIVERSITY

第一章 神经网络概述

- 引言
- 神经网络研究发展简史
- 从人脑神经元到人工神经元
- 人工神经网络构成的基本原理
- 人工神经网络的基本功能
- 小结

2004-8-8
《神经网络导论》——概述
1-1

XIAN JIAOTONG UNIVERSITY

§ 1.1 引言

- 为什么要研究人工神经网络?
- 人工智能与人工神经网络
- 人工神经网络的概念
- 神经网络的研究目的
- 神经网络的研究内容

2004-8-8
《神经网络导论》——概述
1-2

XIAN JIAOTONG UNIVERSITY

§ 1.1 引言

一、为什么要研究人工神经网络?

1. 冯·诺依曼体系结构的计算机发展非常迅速。它可以帮助人类完成许多工作，而且比人更快速、更准确。故称之为“电脑”。它在**数值计算、符号逻辑推理**方面有巨大的优势。

机型	XT	286	386	486	Pent	P II	P III	PIV
主频	4.77M	6M	16M	25M	100M	233M	450M	1.4G
内存	640K	1M	4M	8M	16M	64M	128M	256M
硬盘	10M	20M	40M	200M	1.2G	4G	10G	40G
时间	1983	1984	1986	1989	1993	1997	1999	2000

注：[1] 主频和时间是首次推出时的主频和时间。
 [2] 内存和硬盘是该机型的主流配置。

2004-8-8
《神经网络导论》——概述
1-3

XIAN JIAOTONG UNIVERSITY

§ 1.1 引言

2. 但是很快，人们就发现计算机的能力在某些方面与人相差很远。如视觉信息处理等。可见，这个“电脑”并不健全，它只有**逻辑思维，而缺乏形象思维**。


为什么会出上述情况呢？经过研究发现计算机善于解决（处理）结构性问题，而对非结构性问题则束手无策。

3. 结构性问题和非结构性问题

A. 结构性问题

可以用数学语言清楚而严格地描述，且可将其算法公式化，并映射成计算机程序，然后由计算机逐条地执行该程序的指令，从而得到问题的解。

2004-8-8
《神经网络导论》——概述
1-4



§ 1.1 引言

简单地讲，结构性问题可以程序化计算，也即可以用一个已知的步骤或一组已知的规则来描述。如数值计算、方程求解等。

B. 非结构性问题


如语言或图形识别中，人们无法把自己的认识翻译成严密的机器指令，因而难以用冯·诺伊曼计算机获得满意的结果。

在非结构性问题的解决方面人类比目前的计算机要强多了。

在结构性问题方面，也不一定是计算机强，如算法复杂度极大的TSP（旅行商问题），虽然用计算机可以轻松编程，但要得到问题的解是不现实的。又如，公开密钥技术。

C. 既然Von Neumann计算机对非结构性问题没有有效的解决办法，那么就有必要寻求另外的解决途径。而由于人脑可以轻松的处理这些问题，因此研究人脑是如何解决这些问题的就显得十分重要。

2004-8-8 《神经网络导论》——概述 1-5



§ 1.1 引言

二、人工智能与人工神经网络


1. 人工智能

立足于逻辑运算，符号操作，通过算法来实现智能行为。符号处理方法主要是用计算机模拟人脑的思维功能，重点研究的是机器的思维问题，解决问题的关键在于知识的表示、获取、存储和使用。

符号处理在脑的思维功能模拟等方面取得了很大进展（如深蓝下象棋，各种专家系统等），但对诸如视听觉、联想记忆等问题的处理上往往感到力不从心。

人工智能对高层次（宏观）脑功能的宏观模拟很有效，而对一些低层次的模式处理至今还有一些困难。所以，符号处理不可能全面地解决认知问题和机器智能化问题。

2004-8-8 《神经网络导论》——概述 1-6



§ 1.1 引言


2. 人工神经网络

人工神经网络采用自下而上的方法，从脑的神经系统结构出发来研究脑的功能，研究大量简单神经元的集团信息处理能力。

神经网络对多少年来一直困扰计算机科学和符号处理的一些难题可以得到比较令人满意的解答。

符号处理与神经网络是一种互补的关系。神经网络的研究重点在于模拟和实现人的认知过程中的感知觉过程、形象思维、分布式记忆和自学习自组织过程。而符号处理则侧重于模拟人的逻辑思维。神经网络与符号处理相结合，使人们对人的认知过程有一个较全面的理解。

2004-8-8 《神经网络导论》——概述 1-7



§ 1.1 引言

三、人工神经网络的概念

（人工）神经网络是由**大量简单的基本元件——神经元（Neuron）相互连接**而成的**自适应非线性动态系统**。

单个神经元结构和功能比较简单，大量（ 10^{11} ）神经元组合产生的系统却十分复杂。

系统模型可以用软件描述（数学方程、算法、程序）而最终目标是用硬件实现（半导体、光学器件、分子器件）。

（人工）神经网络反映了人脑功能的若干基本特性，但并非生物系统的逼真描述，只是某种模仿、简化和抽象。

2004-8-8 《神经网络导论》——概述 1-8

XIAN JIAOTONG UNIVERSITY

§ 1.1 引言

四、神经网络的研究目的

神经网络是在许多学科（医学、生物学、生理学、哲学、信息与计算机科学、认知学等）的基础上发展起来的交叉学科。神经网络研究的深入必然带动其它相关学科的发展。各学科研究神经网络的目的不同。

心理学家和认知科学家研究神经网络的目的在于探索人脑加工、储存和搜索信息的机制，弄清人脑功能的机理，建立人类认知过程的微结构理论。

信息处理与计算科学家的研究目的

- a. 解决目前计算机不能解决或不善于解决的大量问题
- b. 构造更加逼近人脑功能的新一代计算机——复杂并行智能信息处理系统——第六代计算机（Neuro-computer神经网络计算机）。

FACETS (Fast Analog Computing With Emergent Transient States)

2004-8-8 《神经网络导论》——概述 1-9

XIAN JIAOTONG UNIVERSITY

§ 1.1 引言

五、神经网络的研究内容

神经网络的研究内容十分广泛，如各种网络模型的构造，以及它们的原理、性能分析，在各个领域的应用等，具体如下：

1. **生物原型研究**：从生理学、心理学、解剖学、脑科学、病理学等生物科学方面。研究神经网络细胞、神经网络、神经系统的生物原型结构及功能机理。
2. **建立理论模型**：根据生物原型的研究，建立神经元，神经网络的理论模型。（概念模型、知识模型、物理化学模型、数学模型等）。

生物原型研究

↓ 科学抽象

建立理论模型

↓ 数学描述

网络模型与算法研究

↓

神经网络应用系统研究

人工神经网络的研究内容

2004-8-8 《神经网络导论》——概述 1-10

XIAN JIAOTONG UNIVERSITY

§ 1.1 引言

理论模型用于神经网络功能和结构分析，可以从定性、定量、静态与动态、微观与宏观等各种观察角度进行研究，是一种广义模型。

3. **网络模型与算法研究**：在理论模型的基础上构作具体的神经网络模型，以实现计算机模拟或准备制作硬件，包括网络学习算法的研究。（技术模型）。
4. **神经网络应用系统研究**：在网络模型与算法研究的基础上，利用人工神经网络组成实际的应用系统。例如完成某种信号处理或模式识别功能。（可以制作专用硬件 / 计算机模拟）。

人工神经网络的研究内容侧重于网络模型与算法和应用系统研究两个方面。

虽然神经网络求解问题不依赖于算法程序，但其结构模型仍可用算法描述，网络结构的研究工作仍借助冯·诺伊曼机，这是当前神经网络研究的特点。

2004-8-8 《神经网络导论》——概述 1-11


XIAN JIAOTONG UNIVERSITY

§ 1.1 引言

这样做虽不能完全显示神经网络的优势，但在论证原理、性能方面却比较方便，同时也可以直接提供某些应用。因为目前硬件实现很难，作为一种过渡，可以对一般的计算机接入一些加速板或利用一些并行处理技术使其更适合于神经网络的运算，直接运用。

从神经网络研究方法看，目前尚未形成统一、完整的理论体系，各种算法性能评价依赖于实验结果；另一方面，神经网络的硬件实现遇到不少困难，在一段时间内，借助计算机插入加速板或采用并行处理技术的过渡方法仍将持续发挥作用。

2004-8-8 《神经网络导论》——概述 1-12




§ 1.2 神经网络研究发展简史

神经网络早期地研究工作可以追溯至二十世纪四十年代。其发展过程大致可以分为以下几个阶段。

1. 奠基时期
 - A. 1943年心理学家W. McCulloch和W. Pitts发表文章，总结了生物神经元的一些基本特性，提出了形式神经元的数学描述与结构方法，即M-P模型，此模型一直沿用至今，可以说他们是神经网络研究的先驱。
 - B. Von Neumann（冯·诺伊曼）。作为指令存储式计算机的奠基人（1945年试制成功），1948年，他在研究工作中比较了人脑结构与存储程序式计算机的根本区别，提出了以简单神经元构成的自再生自动机网络结构。但由于指令存储式计算机技术的发展十分迅速，迫使他放弃了神经网络研究的新途径，继续投身于指令存储式计算机的研究，并做出了巨大贡献。

2004-8-8 《神经网络导论》——概述 1-13




§ 1.2 神经网络研究发展简史

- C. 1949年心理学家Donald Hebb提出神经元之间突触联系强度可变的假设，认为学习过程是在突触上发生的，突触的联系强度随其前后神经元的活动而变化。根据这一假设，Hebb提出了神经元突触的一种具体的学习规律，为神经网络的学习算法奠定了基础。

2. 第一次研究高潮
 - A. 50年代末期，Frank Rosenblatt设计制作了一种多层神经网络“感知机”首次把神经网络的研究从理论探讨付诸工程实践。然后，世界上许多实验室仿效制作感知机，分别应用于文字识别、声音识别以及学习记忆等问题的研究。
 - B. 60年代初，Widrow提出了自适应线性元件网络，简称为Adaline (adaptive linear element)。主要用于雷达天线控制、自适应均衡和回波抵消等方面。后来，在此基础上发展了非线性多层自适应网络。


2004-8-8 《神经网络导论》——概述 1-14



§ 1.2 神经网络研究发展简史

- C. Marvin Minsky 和 Seymour Papert，这次高潮持续了十年左右，不是很久，主要有三个方面原因：
 - ① 数字计算机的发展处于全盛时期，所以人们认为，数字计算机可以解决人工智能、模式识别、专家系统等方面的问题，放松了对神经网络的研究。
 - ② 电子技术工艺水平技术还不高，主要的电子元件是电子管或晶体管，无VLSI。
 - ③ 人工智能学者Minsky 和 Papert 对感知机进行了深入地研究，于1969年出了一本“感知机”的专著。书中论证了：<1> 简单感知机的处理功能有限，不能解决异或问题。<2> 多层网络又很难找到有效的学习算法。
 以上这些原因，使神经网络研究转入低潮，进入萧条时期。

2004-8-8 《神经网络导论》——概述 1-15




§ 1.2 神经网络研究发展简史

3. 萧条期（沉默期）

直至1982年，美国几乎没有什么引人注意的有关神经网络的研究，欧洲和日本受“感知机”一书的影响小。仍有人坚持研究，主要有：

 - A. Kohonen提出的自组织映射模型SOM (selforganizing mapping)。
 - B. Grossberg和Carpenter的自适应谐振理论ART。
 - C. 福岛邦彦 (Fukushima) 的新认知机等。
4. 复兴期（第二次研究高潮）
 - A. John Hopfield, 1982, 1984年，物理学家Hopfield在美国科学院院刊上发表了两篇关于神经网络研究的论文，引起了巨大的反响。这是学术界公认的又一次研究高潮的标志。他的主要贡献有：

2004-8-8 《神经网络导论》——概述 1-16



XI'AN JIAOTONG UNIVERSITY


§ 1.2 神经网络研究发展简史

- ① 引入能量函数 (Lyapunov) 概念, 建立了神经网络稳定性判据。
- ② 利用模拟元件制作神经网络, 为硬件实现奠定了基础。
- ③ 求解了著名的TSP (Travelling Salesman Problem)问题。

B. Rumelhart 和 McClelland 的BP算法。要推动神经网络研究的全面开展, 必须直接解除对多层网络算法的疑虑。1986年, 两人提出了多层网络的“逆推”算法(Back Propagation)BP。该算法从后向前修正各层之间的连接权重, 可以求解感知机所不能解决的问题, 从实践上证明了人工神经网络具有很强的运算能力, 从而否定了Minsky等人的错误结论。

C. 1985年, Hinton, Sejnowsky和Rumelhart提出了Boltzmann机。

2004-8-8
《神经网络导论》——概述
1-17



XI'AN JIAOTONG UNIVERSITY

§ 1.3 从人脑神经元到人工神经元

一、人脑神经元

人的大脑皮层神经元的数目在 $10^{10} - 10^{11}$ 量级, 每立方毫米约有数万个。它的类型很多, 基本结构如图1.1所示。

树突 (dendrite), 又称晶枝, 是细胞体的延伸部, 逐渐变细, 是引入输入信号的突起。

轴突 (axon), 是作为输出端的突起, 只有一个。

突触 (synapse), 树突的全长各部位都可与其它神经元的轴突末梢相互联系, 形成“突触”。

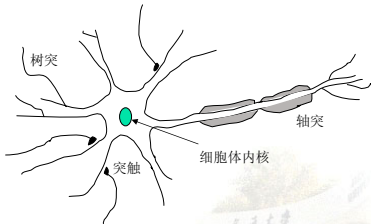



图1.1 神经元结构示意图

2004-8-8
《神经网络导论》——概述
1-18



XI'AN JIAOTONG UNIVERSITY

§ 1.3 从人脑神经元到人工神经元


突触处两神经元并未连通, 它只是发生信息传递的结合部。联系界面间隙约为 $(15 \sim 50) \times 10^{-9}m$ 。突触分为兴奋型和抑制型两种, 相应于神经元之间耦合的极性。每个神经元的突触数目不等, 最高可达 10^5 个。各神经元之间的连接强度和极性有所不同, 且可调整, 基于这一特性, 人脑具有存储信息的功能。

二、人工神经元 (artificial neuron)

基于以上的生物原型, 从不同的观点 (或角度) 出发, 可以抽象出各种人工神经元的模型。

下面介绍一种经典的模型——McCulloch—Pitts模型, 其结构如图1.2所示。

2004-8-8
《神经网络导论》——概述
1-19



XI'AN JIAOTONG UNIVERSITY

§ 1.3 从人脑神经元到人工神经元

对第*i*个神经元, 接受多个其它神经元的信号 x_j , 各突触的连接强度用实系数 w_{ij} 表示, 是对第*j*个输入的加权。所有的输入通过某种运算结合在一起, 称为净输入, 用 Net_i 或 I_i 表示。净输入的表达式很多类型, 简单的如线性加权求和: $Net_i = \sum w_{ij}x_j$ 。这种作用引起神经元的状态变化, 神经元*i*的输出 y_i 是其当前状态的函数。

对第*i*个神经元, 接受多个其它神经元的信号 x_j , 各突触的连接强度用实系数 w_{ij} 表示, 是对第*j*个输入的加权。所有的输入通过某种运算结合在一起, 称为净输入, 用 Net_i 或 I_i 表示。净输入的表达式很多类型, 简单的如线性加权求和: $Net_i = \sum w_{ij}x_j$ 。这种作用引起神经元的状态变化, 神经元*i*的输出 y_i 是其当前状态的函数。

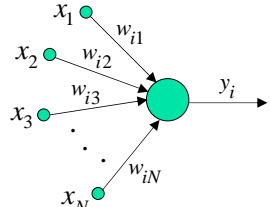


图1.2 McCulloch-Pitts模型示意图

2004-8-8
《神经网络导论》——概述
1-20

XIAN JIAOTONG UNIVERSITY

§ 1.3 从人脑神经元到人工神经元

M-P模型的数学表达式为：

$$y_i = \text{sgn} \left(\sum_j w_{ij} x_j - \theta_i \right) \dots\dots\dots \text{EQ1.1}$$

其中， θ_i 为阈值， $\text{sgn}(\bullet)$ 是符号函数。

如果考虑输出与输入之间的延时，则有

$$y_i(t+1) = \text{sgn} \left(\sum_j w_{ij} x_j(t) - \theta_i \right) \dots\dots\dots \text{EQ1.2}$$

其它各种类型的神经元模型结合需要在以后陆续介绍。

注：以后，省掉人工两字，直接称神经网络。

2004-8-8 《神经网络导论》——概述 1-21

XIAN JIAOTONG UNIVERSITY

§ 1.3 从人脑神经元到人工神经元

三、用大量神经元构成的神经网络的特点

用大量神经元相互连接构成的神经网络将显示出人脑的若干特征。当然这正是我们所希望看到的。

1. 大规模并行处理

人脑神经元之间传递脉冲信号的速度很低，在毫秒量级，比计算机低多了，但因为是大规模并行处理，所以使系统的整体工作极快，在许多问题上可以作出快速判断和处理，速度远高于计算机。人工神经网络的基本结构模仿人脑，具有并行处理的特征，因此可以大大提高工作速度。
2. 分布式存储

人脑的信息存储在突触连接强度上，存储与运算融为一体。分布式存储且冗余很大，每天人脑有大量的神经细胞死亡，不会影响大脑的功能。

2004-8-8 《神经网络导论》——概述 1-22

XIAN JIAOTONG UNIVERSITY

§ 1.3 从人脑神经元到人工神经元

而冯·诺依曼计算机的存储与运算互相独立，知识的存储与数据运算互不相关，需要通过程序沟通，冗余很小。元件的损伤或程序的错误都可能引起严重失常。

3. 自适应（学习）过程

大脑有很强的自适应与自组织特性，后天训练可以开发各种功能。如盲人的听觉、触觉非常灵敏等。

人工神经网络也具有初步的自适应与自组织能力。在学习或训练过程中改变突触权重 W_{ij} ，以适应周围环境的要求。同一神经网络因学习方式及内容不同可具有不同的功能。人工神经网络具有学习能力，可以发展知识。

人工神经网络的学习可以分为有监督学习（Supervised Learning）和无监督学习（Unsupervised Learning）两种方式。

2004-8-8 《神经网络导论》——概述 1-23

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

人工神经网络是典型的并行分布式处理系统。本节首先讨论Rumelhart所列举的研究PDP（Parallel Distributed Processing）模型的八个方面。由于在人工神经网络的研究中通常不必考虑所有八个方面的细节，而常归结为三个方面，所以本节将详细讨论这三个方面。

一、Rumelhart的PDP模型

1986年，Rumelhart列举了研究PDP模型的八个方面，下面结合图1.3对这八个方面进行简要说明。

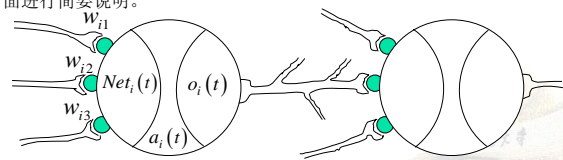


图1.3 Rumelhart的PDP模型

2004-8-8 《神经网络导论》——概述 1-24

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

- 处理单元个数：**网络中包含的处理单元（神经元）的数目，用 N 表示。
- 单元集合的激活 (activation)：**所谓激活是指在指定的时间指定的处理单元所具有的状态，也称**激活状态**。单元 i 在 t 时刻的激活值为 $a_i(t)$ ，系统中所有的处理单元在 t 时刻的激活值用 N 维实向量 $\mathbf{a}(t)$ 表示。激活状态可取连续值，也可取离散值。
- 各单元的**输出函数**：将激活状态 $a_i(t)$ 映射为各单元的输出 $o_i(t)$

$$o_i(t) = f_0[a_i(t)] \dots \dots \dots \text{EQ1.3}$$

输出函数可以有不同规律，如恒同函数、阈值函数、概率统计函数。
- 单元互连模式：**各单元之间相互连接的规律，特定的规律决定了网络的功能特征。通常，**用连接矩阵 \mathbf{W} 来描述互连模式**。

2004-8-8 《神经网络导论》——概述 1-25

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

- 传播规则：**输出矢量 $\mathbf{o}(t)$ 与连接矩阵 \mathbf{W} 给出了各处理单元的输入信号，对各单元综合各输入信号得到净输入 Net_i ，描述此作用的规则称为传播规则。
- 激活规则：**净输入与当前激活状态通过函数 F 的作用共同决定一个新的激活状态，这就是激活规则（状态演变）。对第 i 个处理单元，可表示为
$$a_i(t+1) = F[a_i(t), Net_i(t)] \dots \dots \dots \text{EQ1.4}$$
- 学习规则：**修改单元间的相互连接模式可以改变系统的知识结构。通过样本训练调整**权重 (相互连接强度)**，描述此过程的规则称为学习规则。
- 系统工作环境：**指系统输入信号分布特性的统计规律。

根据神经网络研究的特点，通常主要考虑：（1）神经元功能函数；（2）神经元之间的连接形式；（3）学习（训练）规则等三个方面。

2004-8-8 《神经网络导论》——概述 1-26

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

二、神经元功能函数

神经元功能函数 f ，描述的是神经元在输入信号作用下产生输出信号的规律。是神经元的外特性。

f ：输入信号 $\xrightarrow{\text{(传播规则)}} Net_i(t) \left. \begin{array}{l} \xrightarrow{\text{激活规则}} a_i(t+1) \\ \xrightarrow{F} a_i(t) \end{array} \right\} \xrightarrow{\text{输出函数}} o_i(t+1)$

神经元功能函数包括的过程

f 综合了净输入 Net_i ，激活规则 F 函数和输出函数 f_0 的作用。神经元功能函数有多种形式，利用它们的不同特性可以构成功能各异的神经网络。通常， f 函数可以划分为三种类型：简单的映射关系、动态系统方程或概率统计模型。

简单的映射关系模型 不考虑神经元的时间滞后效应，各神经元构成的输出

2004-8-8 《神经网络导论》——概述 1-27

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

矢量 \mathbf{Y} 与输入矢量 \mathbf{X} 符合某种映射关系。如 $\mathbf{Y} = f(\mathbf{WX} - \boldsymbol{\theta})$ ，其中 $\boldsymbol{\theta}$ 是阈值向量。这种映射可以是线性的或非线性的。

动态系统方程模型 反映了神经元输出与输入之间的延时作用。通常用差分方程或微分方程描述，如 $\mathbf{Y}(t+1) = f(\mathbf{WX}(t) - \boldsymbol{\theta})$ 。

概率统计模型 的输出 \mathbf{Y} 和输入 \mathbf{X} 之间不存在确定性的关系，而是利用一个随机函数说明神经元特性。在此模型中，网络在某一时刻的状态已无意义，而是研究状态的统计规律。

下面介绍一些简单的神经元功能函数实例，主要以差分方程形式给出。

(1) 简单线性模型

$$\mathbf{Y}(t+1) = \mathbf{WX}(t) \dots \dots \dots \text{EQ1.8}$$

或用简单映射关系模型表示为 $\mathbf{Y} = \mathbf{WX}$ 。

2004-8-8 《神经网络导论》——概述 1-28

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

(2) 线性阈值单元 (硬限幅函数)

这是一种非线性模型, 输出只取两种值(0,1)或(+1,-1)。

$$Y(t+1) = \text{sgn}(\mathbf{WX}(t) - \boldsymbol{\theta}) \quad \dots\dots\dots \text{EQ1.9}$$

(3) sigmoid函数模型 (s形函数模型)

输出是限制在两个有限值之间的连续非减函数。

$$Y(t+1) = \tanh\left\{\frac{1}{u_0}[\mathbf{WX}(t) - \boldsymbol{\theta}]\right\} \quad \dots\dots\dots \text{EQ1.10}$$

其中 $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ 是双曲正切函数。 u_0 是控制过零点陡峭程度的参数。

$$\lim_{u_0 \rightarrow 0} \tanh\left(\frac{u}{u_0}\right) = \text{sgn}(u) \quad (\text{如图1.4a所示})$$

2004-8-8 《神经网络导论》——概述 1-29

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

(a) (b)

图1.4 sigmoid函数曲线

2004-8-8 《神经网络导论》——概述 1-30

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

如果想得到0, 1输出, 可将曲线上移, 如图1.4b所示。

$$\frac{1}{2} \left[1 + \tanh\left(\frac{u}{u_0}\right) \right] = \frac{1}{1 + \exp(-2u/u_0)}$$

运算放大器可实现双曲线函数特性。

(4) 多项式阈值函数模型 (高阶关联神经元函数模型或称 $\Sigma - \Pi$ 函数模型)

上面的例子, 神经元净输入作用均为线性函数, 为改善神经元功能, 净输入也可由多项式函数构成, 即各输入信号相乘加权求和。如

$$Net_i = w_{i1}x_1(t) + w_{i2}x_2(t) + w_{i12}x_1(t)x_2(t) \quad \dots\dots\dots \text{EQ1.13}$$

随着输入矢量维数增高, 多项式函数中的交叉相乘项数增多。

2004-8-8 《神经网络导论》——概述 1-31

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

三、神经元之间的连接形式

神经网络是一个复杂的互连系统, 单元之间的互连模式将对网络的性质和功能产生重要的影响, 它决定着网络的知识结构。

1. 前向网络 (前馈网络)

如图1.5, 前向网络可以划分为若干“层”, 各层依次排列, 第*i*层的神经元只接受第*i*-1层神经元给出的信号, 各神经元之间没有反馈。可以用有向无环图来表示。

图1.5 前向网络

2004-8-8 《神经网络导论》——概述 1-32

§ 1.4 人工神经网络构成的基本原理

输入节点无计算功能，以后各层节点均表示有计算功能的神经元，称为计算单元。每个计算单元可有任意个输入，但只有一个输出，它可送到多个神经元做输入（一个轴突，但有多个末梢）。

第一节点层与输出节点层统称为“可见层”，其他中间层称为“隐含层”。

2. 反馈网络

典型的反馈神经网络如图1.6(a)所示，每个节点都表示一个计算单元，同时接受外加输入和其他各节点的反馈输入，每个节点也都直接向外部输出。Hopfield网络就属于这种类型。

在某些网络中，还会包括自反馈；某些反馈型神经网络可以表示为一张完全的无向图，如图1.6(b)。还有许多连接形式，我们就不一一介绍了。

2004-8-8 《神经网络导论》——概述 1-33

§ 1.4 人工神经网络构成的基本原理

图1.6 反馈网络

2004-8-8 《神经网络导论》——概述 1-34

§ 1.4 人工神经网络构成的基本原理

四、学习（训练）

学习功能是神经网络最主要的特征之一。各种学习算法的研究在神经网络理论与实践发展过程中起着重要的作用。当前，神经网络研究的许多课题都致力于学习算法的改进、更新和应用。

学习（训练）过程包括有监督学习和无监督学习两种类型。

1. 有监督学习（Supervised Learning）

所谓有监督学习，是通过外部“示教者”（训练样本/习题）进行学习，对于给定的输入模式，当实际的输出结果和正确的期望输出（导师信号/正确答案）有误差时，网络将通过自动调节机制调节相应的连接强度，使其向减小误差的方向改变，经过重复训练，直到最后使神经网络的输出和正确的结果相符合。

有监督学习的学习过程如图1.7所示。

2004-8-8 《神经网络导论》——概述 1-35

§ 1.4 人工神经网络构成的基本原理

样本训练数据加到网络输入端，同时将相应的期望输出与网络输出相比较得到误差信号，以此控制权重连接强度的调整，直至收敛后给出确定的 W 。当样本情况发生变化时，通过学习可修正 W 值以适应新的环境。

常见的有监督学习算法有Hebb规则和Widrow-Hoff规则（又称 δ 规则、误差修正法等）。

下面简单介绍一下Widrow-Hoff规则。

图1.7 神经网络学习过程示意

2004-8-8 《神经网络导论》——概述 1-36

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

给定样本情况下，随机设置初始权重值。

对第 i 个神经元，设 x_j 为输入矢量的第 j 个元素， w_{ij} 是相应的权值。如果期望输出为 d_i ，而实际输出是 y_i ，那么，权值可以按下式调整。

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(d_i - y_i)x_j \quad \dots\dots\dots \text{EQ1.17}$$

或 $\Delta w_{ij} = w_{ij}(t+1) - w_{ij}(t) = \alpha(d_i - y_i)x_j \quad \dots\dots\dots \text{EQ1.18}$

其中， α 为调整步幅系数，是一个小正数，设 $\varepsilon = (d_i - y_i)$

若 $\varepsilon > 0$ 且 $x_j > 0$ ，为了缩小 ε ，应增加 y_i ，因此应增加 w_{ij} ，即 $\Delta w_{ij} > 0$

若 $\varepsilon > 0$ 且 $x_j < 0$ ，为了缩小 ε ，应增加 y_i ，因此应减小 w_{ij} ，即 $\Delta w_{ij} < 0$

$\varepsilon < 0$ 的情况与此类似。经反复迭代收敛后即可求出合适的 w_{ij} 值。系数 α 的选取对收敛性能有较大影响。

2004-8-8 《神经网络导论》——概述 1-37

XIAN JIAOTONG UNIVERSITY

§ 1.4 人工神经网络构成的基本原理

2. 无监督学习 (Unsupervised Learning)

无监督学习不需要导师信号，学习表现为自适应于输入空间的检测规则，其学习过程为：给系统提供动态输入信号以使各单元以某种方式竞争，“获胜者”神经元本身或其邻域得到增强，其它神经元进一步受到抑制，从而将信号分成有用的多个区域。第四章所要讲的ART网就是这样的。

对于平稳的环境， w 可以达到稳态；若环境发生变化， w 随之改变。

学习（训练）阶段与应用（工作）阶段融为一体，不可分割。这种特征与人脑的学习过程更相似。

2004-8-8 《神经网络导论》——概述 1-38

XIAN JIAOTONG UNIVERSITY

§ 1.5 人工神经网络的基本功能

1. 联想记忆 (Associative Memory, AM)

所谓联想记忆就是通过部分（残缺）线索信息存取或唤起相应的记忆，又称协同记忆。联想记忆可分为自联想（Auto-AM）和异联想（Hetero-AM）。

A. 自联想

设有 M 个样本矢量 $\mathbf{X}^{(s)}$ ，其中 $s = 0, 1, 2, \dots, M-1$ ，

如果输入 $\mathbf{X}' = \mathbf{X}^{(s)} + \Delta$ ， Δ 表示由噪声、干扰或图形缺损等引起的偏差，要求输出 $\mathbf{Y} = \mathbf{X}^{(s)}$ ，即去除噪声或干扰使信号复原。

B. 异联想

设有一一对应的两组样本， $\mathbf{X}^{(s)}$ 和 $\mathbf{Z}^{(s)}$ ，

如果输入有噪声的信号 $\mathbf{X}' = \mathbf{X}^{(s)} + \Delta$ ，要求输出 $\mathbf{Y} = \mathbf{Z}^{(s)}$ 。

2004-8-8 《神经网络导论》——概述 1-39

XIAN JIAOTONG UNIVERSITY

§ 1.5 人工神经网络的基本功能

2. 分类 (classifier)

设输入为 N 维矢量 \mathbf{X} ，共有 M 个类别，用 R_j 表示，其中 $j = 0, 1, 2, \dots, M-1$ ；输出 $y_k, k = 0, 1, 2, \dots, M-1$ 对应于其中的一个类别。它们的关系如图1.8所示。

对 $\mathbf{X} \in R_j$ ，那么

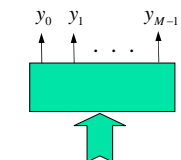
$$y_k = \begin{cases} 1 & \text{当 } k = j \\ 0 & \text{当 } k \neq j \end{cases}$$


图1.8 分类器

3. 优化计算 (优化决策)

将优化约束信息（与目标函数有关）存储于神经网络的连接矩阵 \mathbf{w} 之中，网络的工作状态以动态系统方程式描述。设置一组随机数据作为起始条件，当系统的状态趋于稳定时，网络方程之解作为输出即优化结果。

2004-8-8 《神经网络导论》——概述 1-40

XI'AN JIAOTONG UNIVERSITY

§ 1.5 人工神经网络的基本功能

4. 实现上述功能的人工神经网络应用实例

$y_{AND} = \text{sgn}[x_0 + x_1 - 1]$

$y_{OR} = \text{sgn}[x_0 + x_1 + 1]$

图1.9 神经元实现与、或运算

$y_{XOR} = -\text{sgn}[x_0 x_1]$

图1.10 神经元实现异或运算

2004-8-8 《神经网络导论》——概述 1-41

XI'AN JIAOTONG UNIVERSITY

§ 1.6 小结

- 1.1 引言
 - a. 为什么要研究神经网络？结构性问题与非结构性问题
 - b. 人工智能与神经网络
 - c. 神经网络的概念、研究目的、研究内容。
- 1.2 神经网络发展简史
- 1.3 从人脑神经元到人工神经元

神经网络的特点（大规模并行处理、分布式存储、自适应学习）
- 1.4 神经网络构成的基本原理

Rumelhart的PDP模型；神经元功能函数；连接形式；学习（训练）
- 1.5 神经网络的基本功能

联想记忆（自联想、异联想）；分类；优化计算

2004-8-8 《神经网络导论》——概述 1-42