



人工神经网络算法

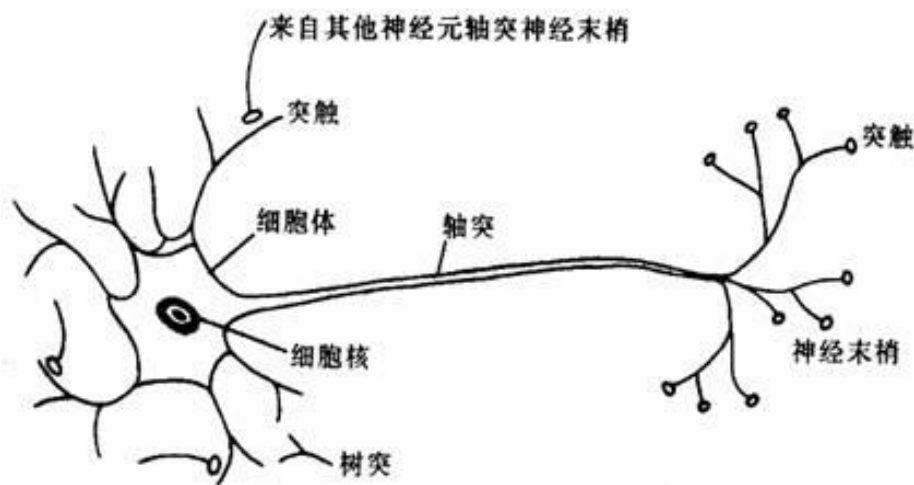
杜逆索

六、人工神经网络算法

1. 基本概念

1.1 生物神经元模型

人的神经系统是由众多神经元相互连接而成的一个复杂系统，神经元又叫神经细胞，它是神经组织的基本单位。

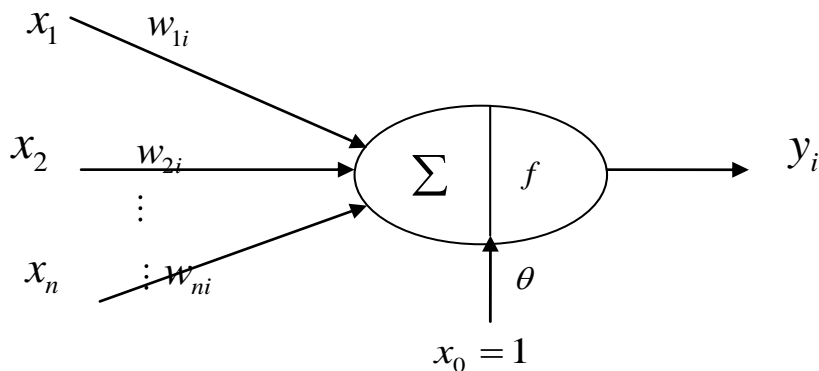


六、人工神经网络算法

1. 基本概念

1.2 人工神经元模型

人工神经元模型就是用人工方法模拟生物神经元而形成的模型,是对生物神经元的抽象、模拟与简化,它是一个多输入、单输出的非线性元件,单个神经元是前向型的。





六、人工神经网络算法

1. 基本概念

1.2 人工神经元模型

在神经元中,对信号进行处理采用的是数学函数,通常称为激活函数、激励函数或挤压函数,其输入、输出关系可描述为

$$\begin{cases} u_j = f\left(\sum_{i=1}^n w_{ij}x_i - \theta_j\right) \\ y = f(u_j) \end{cases}$$

式中 $x_i (i=1,2,\dots,n)$ 是从其它神经元传来的输入信号; θ_j 是该神经元的阈值; w_{ij} 表示从神经元 i 到神经元 j 的连接权值; $f(\square)$ 为激活函数或挤压函数。



六、人工神经网络算法

1. 基本概念

1.2 人工神经元模型

神经元采用了不同的激活函数,使得神经元具有不同的信息处理特性,并且神经元的处理特性是决定神经网络整体性能的主要因素之一。

下面介绍四种常用的激活函数形式:

(1) 阈值型函数,即 $f(x)$ 为阶跃函数。

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

具有这一作用方式的神经元称为阈值型神经元,是神经元模型中最简单的一种,经典的M-P模型神经元就属于这一类。



六、人工神经网络算法

1. 基本概念

1.2 人工神经元模型

神经元采用了不同的激活函数,使得神经元具有不同的信息处理特性,并且神经元的处理特性是决定神经网络整体性能的主要因素之一。

下面介绍四种常用的激活函数形式:

(2) 分段线性函数。此函数特点是:神经元的输入与输出在一定区间内满足线性关系。这类函数也称为伪线性函数,表达式如下:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x \leq x_c \\ 1, & x_c < x \end{cases}$$



六、人工神经网络算法

1. 基本概念

1.2 人工神经元模型

神经元采用了不同的激活函数,使得神经元具有不同的信息处理特性,并且神经元的处理特性是决定神经网络整体性能的主要因素之一。

下面介绍四种常用的激活函数形式:

(3) Sigmoid函数。Sigmoid函数也叫S型函数,通常是在(0,1)或(-1,1)内连续取值的单调可微分函数,它是一类非常重要的激活函数,无论神经网络用于分类、函数逼近或优化,Sigmoid函数都是常用的激活函数。常用指数或正切等一类曲线表示:

$$f(x) = \frac{1}{1 + e^{-\lambda x}} \quad \text{或} \quad f(x) = \frac{1 - e^{-\lambda x}}{1 + e^{-\lambda x}}$$

其中 λ 又称为Sigmoid函数的增益,其值决定了函数非饱和段的斜率, λ 越大,曲线越陡。



六、人工神经网络算法

1. 基本概念

1.2 人工神经元模型

神经元采用了不同的激活函数,使得神经元具有不同的信息处理特性,并且神经元的处理特性是决定神经网络整体性能的主要因素之一。

下面介绍四种常用的激活函数形式:

(4) 高斯函数。高斯函数(也称钟型函数)也是极为重要的一类激活函数,常用于径向基神经网络(RBF网络),其表达式为:

$$f(x) = e^{-\frac{x^2}{\delta^2}}$$

式中 δ 称为高斯函数的宽度或扩展常数。 δ 越大,函数曲线就越平坦;反之, δ 越小,函数曲线就越陡峭。



六、人工神经网络算法

1. 基本概念

1.3 主要的神经网络模型

目前使用的比较典型的一些神经网络模型主要有以下几类:

1. 误差后向传播(Back Propagation , BP)神经网络

该类神经网络模型是一种多层映射神经网络,采用的是最小均方差的学习方式。

2. 径向基函数(Radial Basis Function , RBF)神经网络

RBF网络采用3层结构,层间的神经元的连接方式同BP网络类似,也是采用层间全连接、层内无连接方式。



六、人工神经网络算法

1. 基本概念

1.3 主要的神经网络模型

目前使用的比较典型的一些神经网络模型主要有以下几类:

3. Hopfield 网络

Hopfield 网络由相同的神经元构成的单层、并且不具学习功能的自联想网络。网络的权值按一定规则计算出来,一经确定就不再改变,而网络中各神经元的状态在运行过程中不断更新,网络演变到稳态时各神经元的状态便是问题之解。

Hopfield 网络有离散型和连续型两种。



六、人工神经网络算法

1. 基本概念

1.3 主要的神经网络模型

目前使用的比较典型的一些神经网络模型主要有以下几类:

4. 随机型神经网络

随机型神经网络其基本思想是:不但让网络的误差和能量函数向减小的方向变化,而且还可按某种方式向增大的方向变化,目的是使网络有可能跳出局部极小值而向全局最小点收敛。随机型神经网络的典型算法是模拟退火算法。



六、人工神经网络算法

1. 基本概念

1.3 主要的神经网络模型

目前使用的比较典型的一些神经网络模型主要有以下几类:

5. Kohonen 网络

Kohonen 网络含有两层,输入缓冲层用于接收输入模式,输出层的神经元一般按正则二维阵列排列,每个输出神经元连接至所有输入神经元, Kohonen 网络是一种以无导师方式进行网络训练的网络,它通过自身训练,自动对输入模式进行分类。



六、人工神经网络算法

1. 基本概念

1.3 主要的神经网络模型

目前使用的比较典型的一些神经网络模型主要有以下几类:

6. 玻耳兹曼机

玻耳兹曼机是由Hinton等人提出的, 它的思想主要来源于统计物理学。它在本质上是建立在Hopfield网络基础上的, 具有学习能力, 能够通过一个模拟退火过程寻求解答。



六、人工神经网络算法

1. 基本概念

1.3 主要的神经网络模型

目前使用的比较典型的一些神经网络模型主要有以下几类:

7. 对向传播神经网络

对向传播(Counter-Propagation)神经网络(CP网络)将Kohonen特征映射网络与Grossberg基本竞争型网络巧妙结合的一种新型特征映射网络。



六、人工神经网络算法

2.BP算法的原理

2.1 Delta学习规则的基本原理

Delta学习规则又称梯度法或最速下降法,其要点是改变单元间的连接权重来减小系统实际输出与期望输出间的误差。

假定神经元权值修正的目标是极小化标量函数 $F(w)$,其中 $F()$ 函数代表神经网络的误差目标函数, w 代表神经网络中的某一连接权值,如果神经元的当前权值为 $w(t)$,并且假设下一时刻的权值调节公式为

$$w(t+1) = w(t) + \Delta w(t)$$

式中, $\Delta w(t)$ 表示当前时刻的权值修正方向。显然,我们期望每次权值修正都能满足

$$F(w(t+1)) < F(w(t))$$

这样神经网络输出的误差才会随着训练过程的进行而不断地向最小化的目标靠近。对 $F(w(t+1))$ 进行一阶泰勒公式展开,得

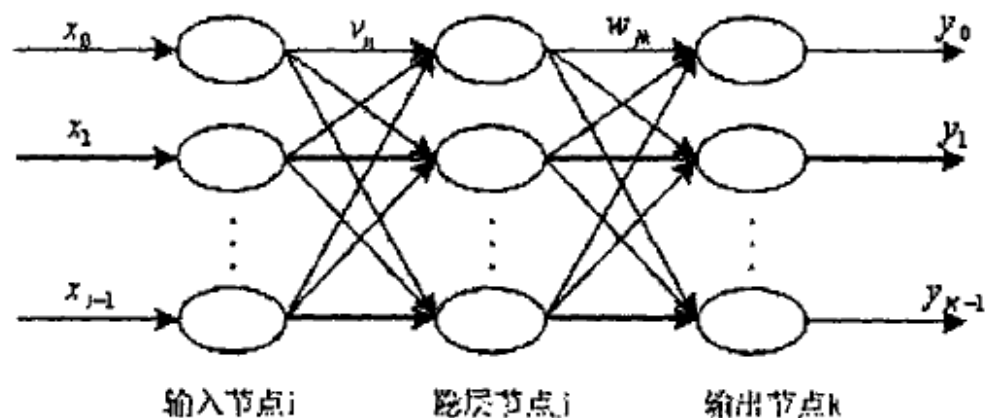
$$F(w(t+1)) = F(w(t) + \Delta w(t)) \approx F(w(t) + g^T(t) \Delta w(t))$$

六、人工神经网络算法

2.BP算法的原理

2.2 BP神经网络的结构

BP神经网络是具有三层或三层以上的阶层型神经网络,由输入层、隐含层和输出层构成,相邻层之间的神经元全互连,同一层内的神经元无连接。





六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

BP算法的主要思想是从后向前(反向)逐层传播输出层的误差,以间接算出隐层误差。

算法分为两个阶段:

第一阶段(正向传播过程)输入信息从输入层经隐层逐层计算各单元的输出值。

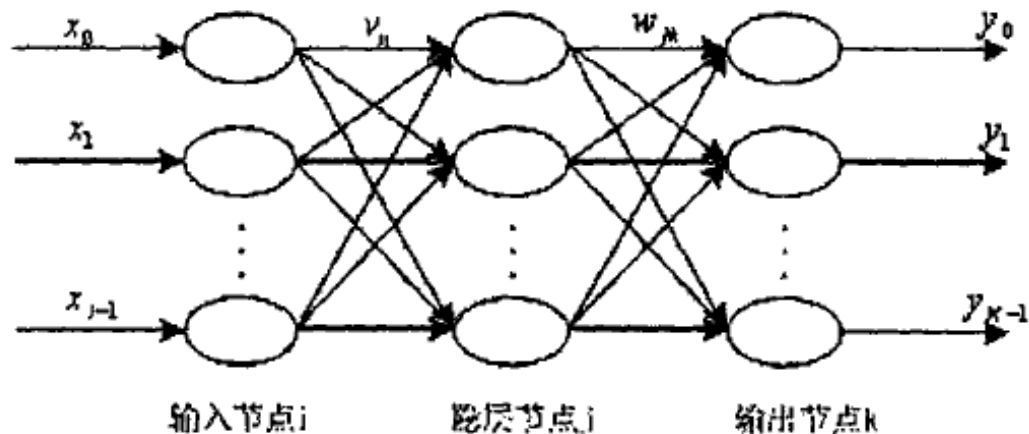
第二阶段(反向传播过程)输出误差逐层向前算出隐层各单元的误差,并用此误差修正前层权值。

六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

1.正向计算输出阶段



以上图的BP神经网络为例，假设BP网络的输入层节点数为 I ，隐含层节点数为 J ，输出层节点数为 K 。输入向量为 $X^p = (x_0, x_1, x_2, \dots, x_{I-1})$ ，其期望输出向量为 $D^p = (d_0, d_1, d_2, \dots, d_{K-1})$ ，则有：



六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

1.正向计算输出阶段

(1) 输入层: $O_i = x_i, i = 0, 1, 2, \dots, I-1$

(2) 隐含层: 为简化推导,把各点的阈值当作一种特殊的连接权值,其对应的输入恒为-1。对于第j个神经元的输入为: $net_j = \sum_{i=0}^I v_{ij} O_i$

, 其中 $O_I = -1$ 、 v_{Ij} 为阈值;

其第j个节点的输出为: $O_j = f(net_j)$, $j = 0, 1, 2, \dots, J-1$



六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

1.正向计算输出阶段

(3) 输出层：同理，对于第 k 个神经元的输入为

$$net_k = \sum_{j=0}^J w_{jk} O_j$$

，其中 $O_j = -1$ 、 w_{jk} 为阈值；

其第 k 个节点的输出为： $O_k = f(net_k)$ ， $k = 0, 1, 2, \dots, K-1$

定义BP网络的能量函数(误差函数)为

$$E_p = \frac{1}{2} \sum_{k=0}^{K-1} (d_k^p - O_k^p)^2$$

式中： E_p —— P 的输出误差；

d_k^p ——样本 P 的期望输出；

O_k^p ——输出层神经元的实际输出。



六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

2.误差反向传播阶段(网络权值修正阶段)

通过调整权值和阈值,使得误差能量达到最小时,网络趋于稳定状态,学习结束。

(1) 输出层与隐含层之间的权值调整。对每一个 w_{jk} 的修正值为:

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \frac{\partial E}{\partial net_k} \cdot \frac{\partial net_k}{\partial w_{jk}}$$

式中: η 为学习步长,取值介于(0,1), 对式 $net_k = \sum_{j=0}^J w_{jk} O_j$ 求偏导得:

$$\frac{\partial net_k}{\partial w_{jk}} = O_j \quad \text{记:} \quad \delta_k = -\frac{\partial E}{\partial net_k}$$



六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

2.误差反向传播阶段(网络权值修正阶段)

有:
$$\delta_k = -\frac{\partial E}{\partial net_k} = -\frac{\partial E}{\partial O_k} \cdot \frac{\partial O_k}{\partial net_k} = (d_k - O_k) f'(net_k)$$

将式子整理得
$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\frac{\partial E}{\partial net_k} \cdot \frac{\partial net_k}{\partial w_{jk}} = \eta \delta_k O_j$$

六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

2.误差反向传播阶段(网络权值修正阶段)

(2) 隐含层与输入层的权值调整。同理,对每一个 v_{ij} 的调整值为:

$$\Delta v_{ij} = -\eta \frac{\partial E}{\partial v_{ij}} = -\eta \frac{\partial E}{\partial net_j} \cdot \frac{\partial net_j}{\partial v_{ij}} = \eta \left(-\frac{\partial E}{\partial net_j} \right) O_i = \eta \delta_j O_i$$

其中,

$$\delta_j = -\frac{\partial E}{\partial net_j} = -\frac{\partial E}{\partial O_j} \cdot \frac{\partial O_j}{\partial net_j}$$

$$\text{再根据: } -\frac{\partial E}{\partial O_j} = -\frac{\partial}{\partial O_j} \left(\frac{1}{2} \sum_{k=0}^{K-1} (d_k - O_k)^2 \right) = \sum_{k=0}^{K-1} (d_k - O_k) \frac{\partial O_k}{\partial net_k} \cdot \frac{\partial net_k}{\partial O_j}$$

$$= \sum_{k=0}^{K-1} (d_k - O_k) f'(net_k) w_{jk} = \sum_{k=0}^{K-1} \delta_k w_{jk}$$



六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

2.误差反向传播阶段(网络权值修正阶段)

得:
$$\delta_j = f'(net_j) \sum_{k=0}^{K-1} \delta_k w_{jk}$$

综上所述,若BP神经网络每一层的激活函数均取单极S型函数,即

$$f(net) = \frac{1}{1 + e^{-net}}$$



六、人工神经网络算法

2.BP算法的原理

2.3 BP神经网络的算法描述

2.误差反向传播阶段(网络权值修正阶段)

则可以方便地计算出该网络各层的权值修正量:

(1) 对于输出层:

$$\Delta w_{jk} = \eta O_j (d_k - O_k) f'(net_k) = \eta O_j (d_k - O_k) O_k (1 - O_k)$$

(2) 对于隐含层:

$$\Delta v_{ij} = \eta \cdot f'(net_j) \sum_{k=0}^{K-1} \delta_k w_{jk} \cdot O_i = \eta O_j (1 - O_j) \sum_{k=0}^{K-1} \delta_k w_{jk} \cdot O_i$$



六、人工神经网络算法

2.BP算法的原理

2.4 标准BP神经网络的工作过程

BP神经网络的总体步骤大致如下所示:

(1) 权值初始化: $w_{ij} = \text{Random}(\bullet)$ 、 $w_{jk} = \text{Random}(\bullet)$, 其中 w_{ij} 表示网络输入层单元到隐含层单元的连接权值; w_{jk} 表示网络隐含层单元到输出层单元的连接权值。

(2) 依次输入P个学习样本。设当前输入为第p个样本。

(3) 依次计算各层的输出: O_j 、 O_k ; 其中 O_j 为隐含层上第j个神经元的输出。 O_k 为输出层上第k个神经元的输出。



六、人工神经网络算法

2.BP算法的原理

2.4 标准BP神经网络的工作过程

(4) 根据式 $\delta_k = -\frac{\partial E}{\partial net_k} = -\frac{\partial E}{\partial O_k} \cdot \frac{\partial O_k}{\partial net_k} = (d_k - O_k) f'(net_k)$

和 $\delta_j = f'(net_j) \sum_{k=0}^{K-1} \delta_k w_{jk}$,求网络各层的反传误差:

$$\delta_k = (d_k - O_k) f'(net_k)$$

$$\delta_j = f'(net_j) \sum_{k=0}^{K-1} \delta_k w_{jk}$$

并记下各个 $O_j^{(p)}$ 、 $O_k^{(p)}$ 的值;



六、人工神经网络算法

2.BP算法的原理

2.4 标准BP神经网络的工作过程

(5) 记录已经学习过的样本个数 p 。如果 $p < P$,转到步骤 (2) 继续计算;如果 $p = P$,转到步骤 (6) 。

(6) 按权值修正公式修正各层的权值和阈值。

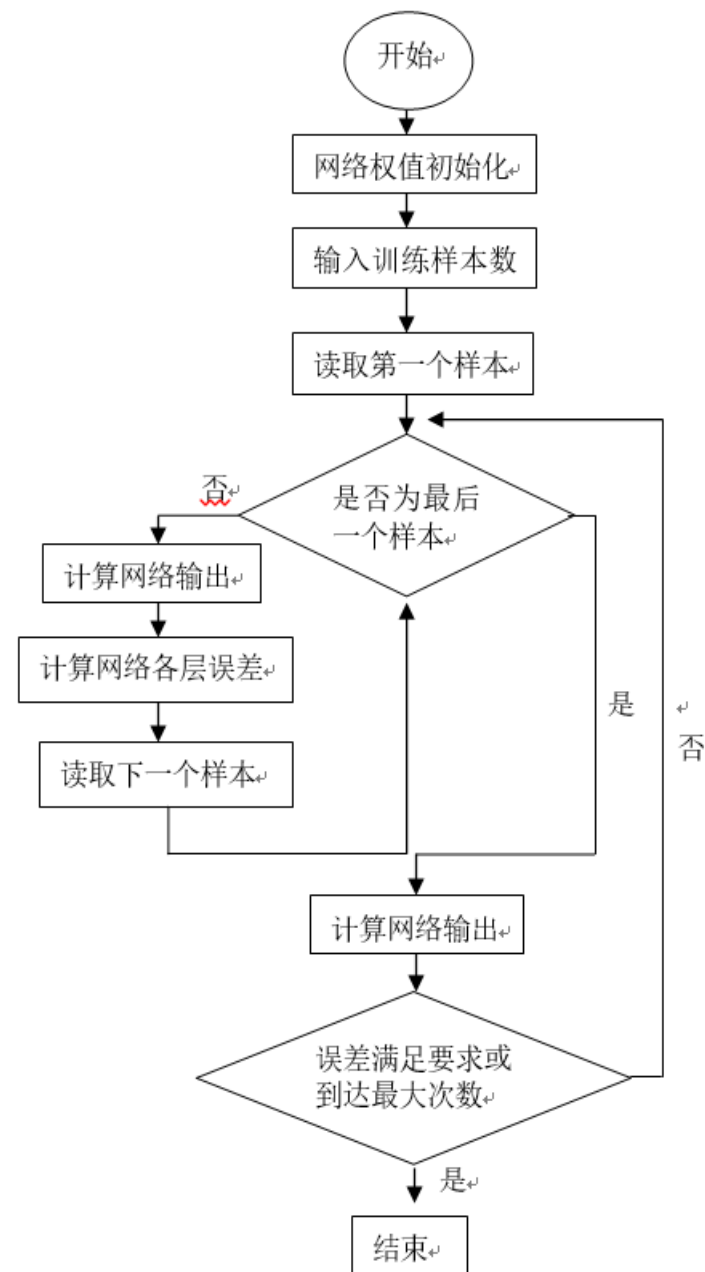
(7) 按新的权值再计算 $O_j^{(p)}$ 、 $O_k^{(p)}$ 和 $E_{\text{总}} = \frac{1}{2N} \sum_{p=0}^{P-1} \sum_{k=0}^{K-1} (d_k^p - O_k^p)^2$ ，若对

每个 p 样本和相应的第 k 个输出神经元,都满足 $|d_k^{(p)} - O_k^{(p)}| < \varepsilon$,或达到最大学习次数,则终止学习。否则转到步骤 (2) 继续新一轮的网络学习。

六、人工神经网络算法

2.BP算法的原理

2.4 标准BP神经网络的工作过程





六、人工神经网络算法

3. BP算法的特点及应用

3.1 BP算法的特点

优点：

- 1 非线性映射能力
- 2 自学习和自适应能力
- 3 泛化能力
- 4 容错能力

缺点：

1. 局部极小化问题
2. 收敛速度慢
3. 结构选择不一
4. 应用实例与网络规模的矛盾问题
5. 预测能力和训练能力的矛盾问题
6. 样本依赖性问题



六、人工神经网络算法

3. BP算法的特点及应用

3.2 BP算法的应用神经网络

目前主要应用于以下几个领域:

- 1.信息领域:用于信号处理、模式识别、数据压缩等方面。
- 2.自动化领域:在系统辨识、神经控制器、智能检测等方面取得了发展。
- 3.工程领域:如汽车工程、军事工程、化学工程、水利工程等。
- 4.医学领域:可用于检测数据分析、生物活性研究、医学专家系统等。
- 5.经济领域:用于识别信贷客户、股票预测和证券市场分析等方面。



六、人工神经网络算法

4. 小结

本章首先介绍了生物与人工神经元模型，然后简单列举了主要的神经网络模型，重点剖析了BP算法，阐述了它的原理、结构、工作过程，最后介绍了BP算法的特点以及应用。