# [BP人工神经网络简](http://www.cnblogs.com/luxiaoxun/archive/2012/12/10/2811309.html)介

## 一、神经网络概念与适合领域

神经网络最早的研究是 40 年代心理学家 Mcculloch 和数学家 Pitts 合作提出的 ，他们提出的MP模型拉开了神经网络研究的序幕。

神经网络的发展大致经过 3 个阶段：1947～1969 年为初期，在这期间科学家们提出了许多神经元模型和学习规则，如 MP 模型、HEBB 学习规则和感知器等；60 年代末期至 80 年代中期，神经网络控制与整个神经网络研究一样，处于低潮。在此期间，科学家们做了大量的工作，如 Hopfield 教授对网络引入能量函数的概念，给出了网络的稳定性判据，提出了用于联想记忆和优化计算的途径。1984年，Hiton 教授提出 Bol tzman 机模型；1986年 Kumelhart 等人提出误差反向传播神经网络，简称 BP 网络。目前，BP网络已成为广泛使用的网络。1987年至今为发展期，在此期间，神经网络受到国际重视，各个国家都展开研究，形成神经网络发展的另一个高潮。

人工神经网络（ANN）受到生物学的启发是生物神经网络的一种模拟和近似，它从结构、实现机理和功能上模拟生物神经网络。从系统观点看，人工神经元网络是由大量神经元通过极其丰富和完善的连接而构成的自适应非线性动态系统。人工神经网络，因为生物的学习系统是由相互连接的神经元组成的异常复杂的网络，其中每一个神经元单元有一定数量的实值输入，并产生单一的实数值输出。1960 年威德罗和霍夫率先把神经网络用于自动控制研究。神经网络以其独特的结构和处理信息的方法，在许多实际应用领域中取得了显著的成效，主要应用如下：自动控制领域、处理组合优化问题、模式识别、图像处理、传感器信号处理、机器人控制、信号处理、卫生保健、医疗、经济、化工领 域、焊接领域、地理领域、数据挖掘、电力系统、交通、军事、矿业、农业和气象等领域。

## 二、神经网络基本结构

人工神经网络由神经元模型构成，这种由许多神经元组成的信息处理网络具有并行分布结构。每个神经元具有单一输出，并且能够与其它神经元连接；存在许多（多重）输出连接方法，每种连接方法对应一个连接权系数。可把 ANN 看成是以处理单元 PE(processing element) 为节点，用加权有向弧(链)相互连接而成的有向图。令来自其它处理单元(神经元)i的信息为Xi，它们与本处理单元的互相作用强度为 Wi，i=0,1,…，n-1，处理单元的内部阈值为 θ。那么本神经元的输入为：

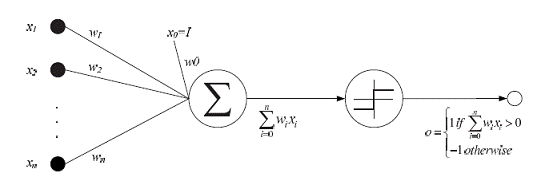
IMG_256

而处理单元的输出为：

IMG_257

式中，xi为第 i 个元素的输入，wi 为第 i 个元素与本处理单元的互联权重。f 称为激发函数(activation function)或作用函数。它决定节点(神经元)的输出。该输出为 1 或 0 取决于其输入之和大于或小于内部阈值 θ。

下图所示神经元单元由多个输入Xi，i=1,2,...,n和一个输出y组成。中间状态由输入信号的权和表示，而输出为：



## 三、训练网络

神经网络结构被设计完成，有了输入、输出参数后，我们就要对网络进行训练。神经网络的训练有包括感知器训练、delta 规则训练和反向传播算法等训练，其中感知器训练是基础。

## 四、感知器和 delta 训练规则

理解神经网络第一步是从对抽象生物神经开始,本文用到的人工神经网络系统是以被称为感知器的单元为基础，如图所示。感知器以一个实数值向量作为输入，计算这些输入的线性组合，如果结果大于某个阈值，就输出 1，否则输出 -1，如果 x 从 1 到 n，则感知器计算公式如下：

IMG_259

其中每个 wi 是一个实数常量，或叫做权值，用来决定输入 xi 对感知器输出的贡献率。特别地，-w0是阈值。

尽管当训练样例线性可分时，感知器法则可以成功地找到一个权向量，但如果样例不是线性可分时它将不能收敛，因此人们设计了另一个训练法则来克服这个不足，这个训练规则叫做 delta 规则。感知器训练规则是基于这样一种思路--权系数的调整是由目标和输出的差分方程表达式决定。而 delta 规则是基于梯度降落这样一种思路。这个复杂的数学概念可以举个简单的例子来表示。从给定的几点来看，向南的那条路径比向东那条更陡些。向东就像从悬崖上掉下来，但是向南就是沿着一个略微倾斜的斜坡下来，向西象登一座陡峭的山，而北边则到了平地，只要慢慢的闲逛就可以了。所以您要寻找的是到达平地的所有路径中将陡峭的总和减少到最小的路径。在权系数的调整中，神经网络将会找到一种将误差减少到最小的权系数的分配方式。这部分我们不做详细介绍，如有需要大家可参考相关的人工智能书籍。

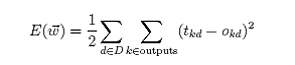
## 五、反向传播算法

人工神经网络学习为学习实数值和向量值函数提供了一种实际的方法，对于连续的和离散的属性都可以使用。并且对训练数据中的噪声具有很好的健壮性。反向传播算法是最常见的网络学习算法。这是我们所知用来训练神经网络很普遍的方法，反向传播算法是一种具有很强学习能力的系统，结构比较简单，且易于编程。

鲁梅尔哈特(Rumelhart)和麦克莱兰(Meclelland)于 1985 年发展了 BP 网络学习算法，实现了明斯基的多层网络设想。BP网络不仅含有输入节点和输出节点，而且含有一层或多层隐(层)节点。输入信号先向前传递到隐藏节点，经过作用后，再把隐藏节点的输出信息传递到输出节点，最后给出输出结果。节点的激发函数一般选用 S 型函数。

反向传播(back－propagation，BP)算法是一种计算单个权值变化引起网络性能变化值的较为简单的方法。由于BP算法过程包含从输出节点开始，反向地向第一隐含层(即最接近输入层的隐含层)传播由总误差引起的权值修正，所以称为"反向传播"。反向传播特性与所求解问题的性质和所作细节选择有极为密切的关系。

对于由一系列确定的单元互连形成的多层网络，反向传播算法可用来学习这个多层网络的权值。它采用梯度下降方法试图最小化网络输出值和目标值之间的误差平方，因为我们要考虑多个输出单元的网络，而不是像以前只考虑单个单元，所以我们要重新计算误差E，以便对所有网络输出的误差求和:



Outpus 是网络输出单元的集合，tkd 和 okd 是与训练样例 d 和第 k 个输出单元的相关输出值.

反向传播算法的一个迷人特性是：它能够在网络内部的隐藏层发现有用的中间表示:

1.训练样例仅包含网络输入和输出，权值调节的过程可以自由地设置权值，来定义任何隐藏单元表示，这些隐藏单元表示在使误差E达到最小时最有效。

2.引导反向传播算法定义新的隐藏层特征，这些特征在输入中没有明确表示出来，但能捕捉输入实例中与学习目标函数最相关的特征

反向传播训练神经元的算法如下：

