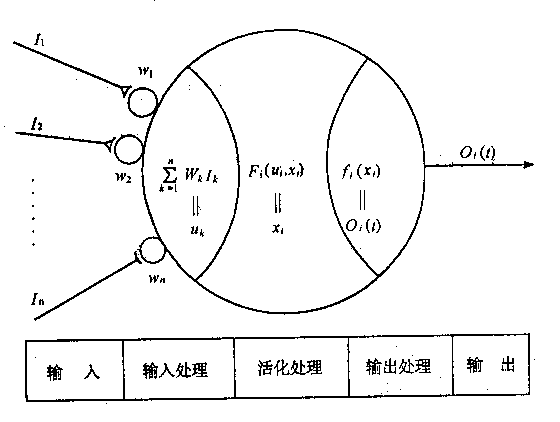
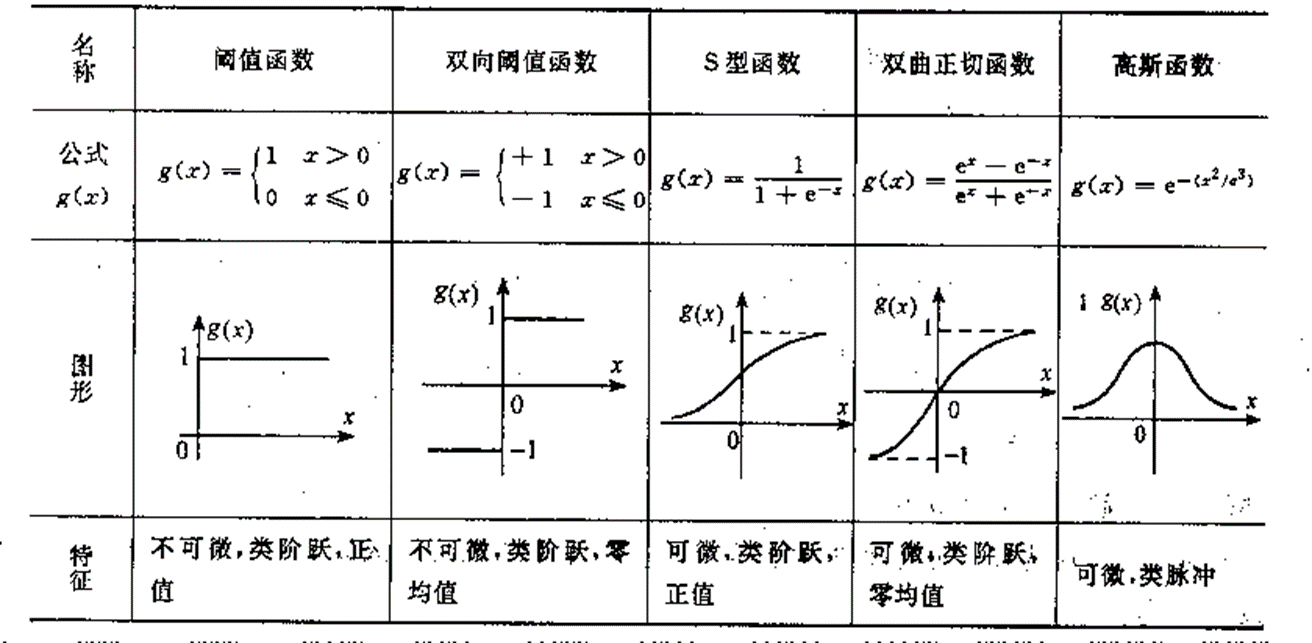
神经网络复习重点总结

### 神经网络基本原理

**神经元模型：**



一般网络输出不做处理，直接作为下一层的输入。其中活化函数一般有以下非线性函数：



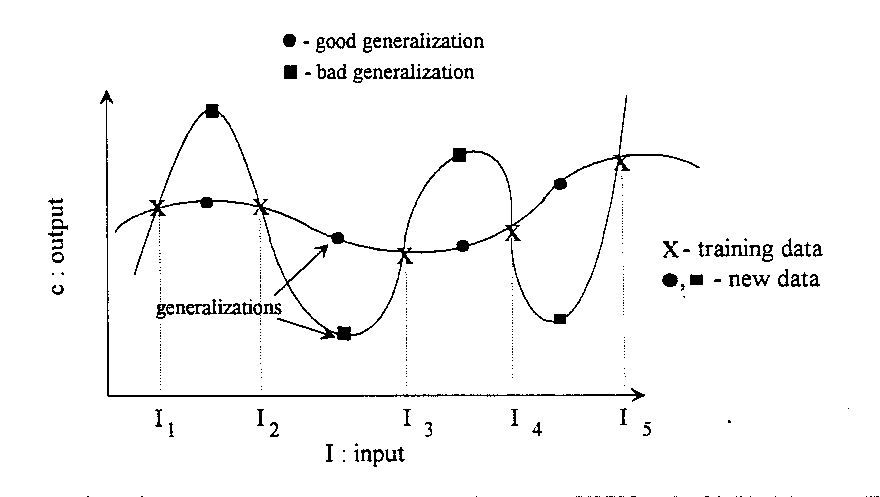
**神经网络一般构建过程：**

1.训练与学习过程：利用训练集对网络进行训练学习

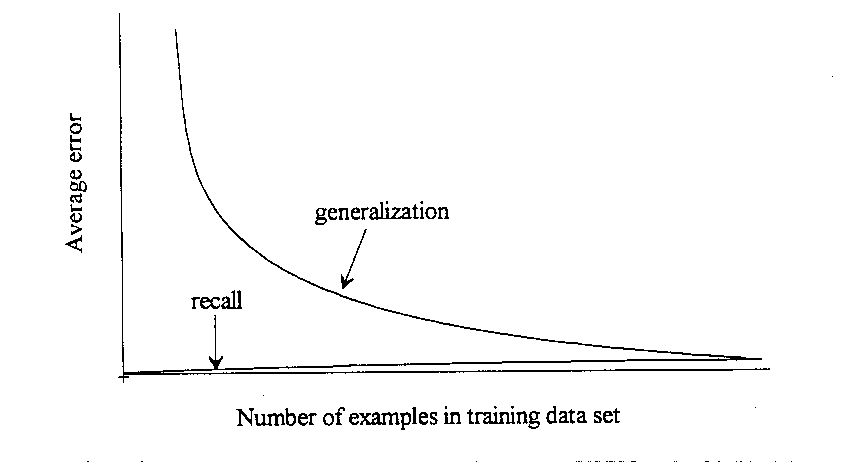
2.回想过程：对训练集的数据重新测试，验证训练效果

3.泛化过程：对测试集的数据进行测试，验证训练效果

好的网络要求回想误差与泛化误差均较小，且差异相当。



训练良好的网络误差曲线：



可能导致网络训练效果不好的原因：

1.数据集选取不当，没有代表性。

2.过度训练导致过拟合，表现为回想误差与泛化误差差距较大。

4.网络模型或算法不合理。

5.超参数（学习速率，学习精度等）设置不当。

6.训练步骤有问题（如归一化时未带入测试集数据）。

**神经网络的性质：**

1.神经网络的优点：

1. 信息存储在每一个节点中。
2. 神经网络具有学习能力。
3. 神经网络允许广泛的知识索引。
4. 神经网络更适合处理嘈杂，不完整或不一致的数据。
5. 神经网络模仿人类学习过程。
6. 自动抽象。
7. 可在线使用的潜力

2. 神经网络与经验模型的比较：

1. 由于微观特征的概念，神经网络比经验模型具有更好的过滤能力。
2. 神经网络由于具有特定的训练算法而比经验模型更具适应性。
3. 神经网络是真正的多入多出系统。

3.神经网络的局限性：

1. 需要较长的训练时间。
2. 需要大量的训练数据。
3. 不能保证结果最佳。
4. 不能保证百分之百可靠。

4.神经网络的潜在应用：

分类问题、预测和优化、过程预测，监测，诊断，建模和控制、数据过滤、专家系统等。

### MP Model ( McCulloch and Pitts in 1943)

**MP模型概念：**将神经元视为二值开关元件按不同方式组合可以完成各种逻辑运算，将这种逻辑神经元模型成为MP模型。（MP模型不具备学习能力，权值恒定，没有学习过程）

**MP模型的创新之处：**1.给出了神经元的形式化数学描述和网络结构方法。2.从数学上证明了单个神经元能执行逻辑运算，从而开创了人工神经网络研究时代。

**MP模型结构：**

**E**

**E**

**I**

**I**

Y{0,1}

**输入条件 输出**

\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_

∑E - ∑I≥T Y=1

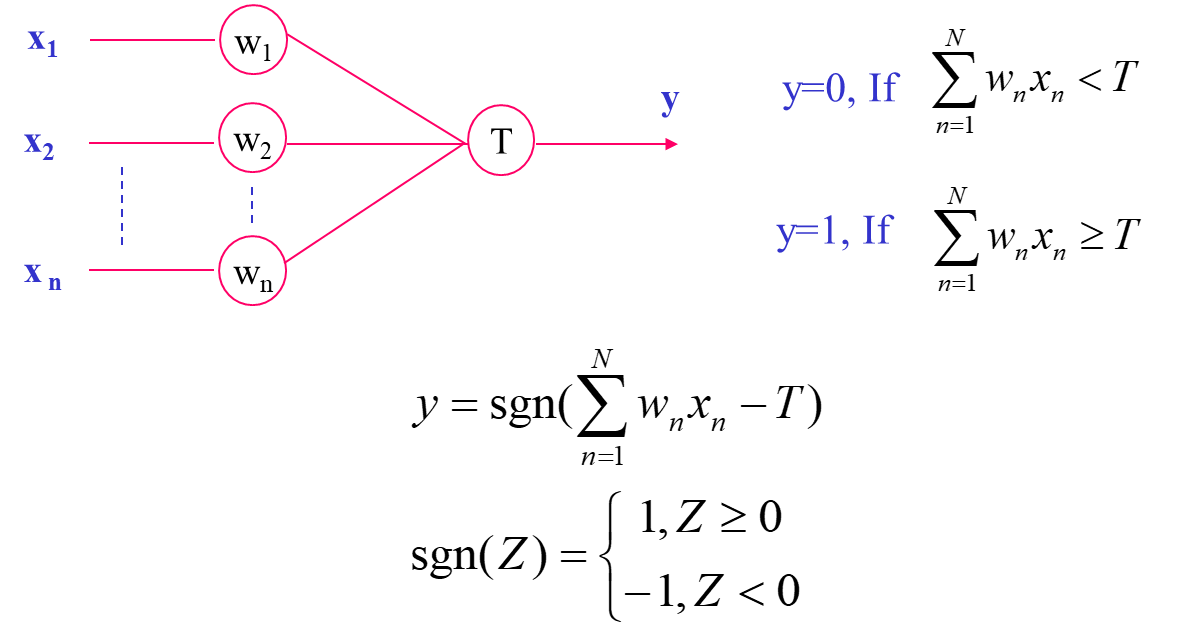
∑E - ∑I<T Y=0

E：激励作用

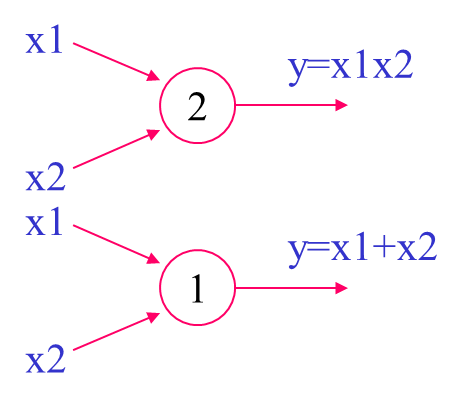
I：抑制作用

T：阈值

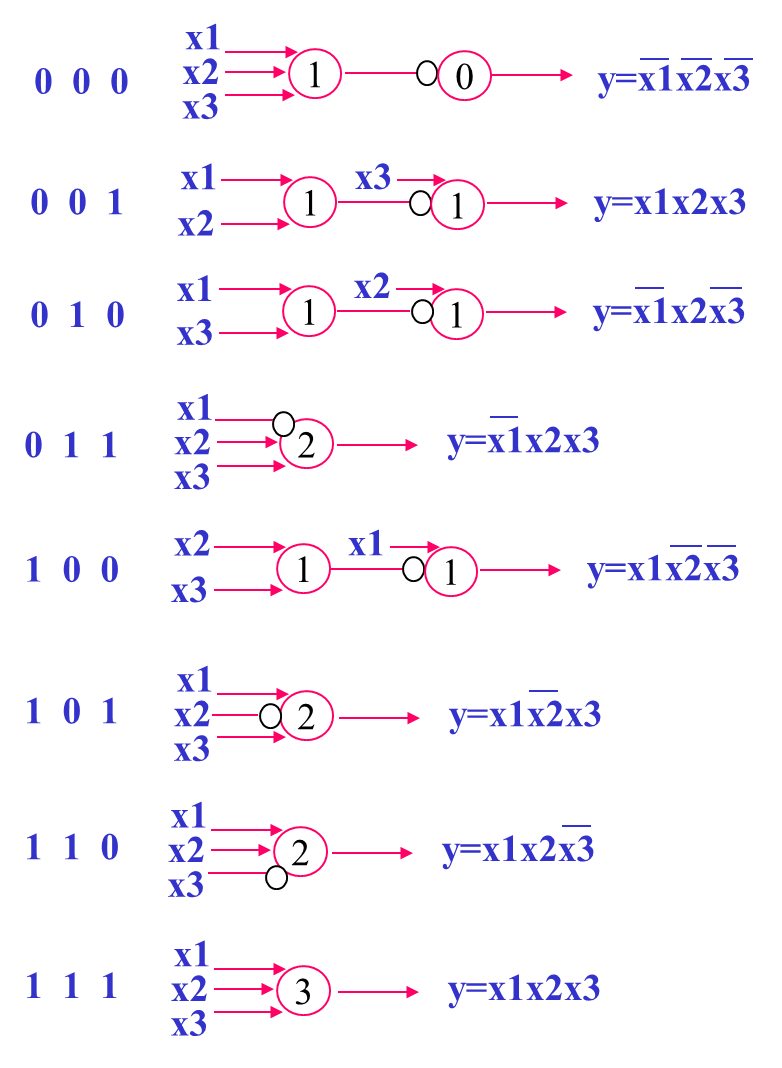
**典型MP模型举例：（Wi取值为正或负，即激励作用或抑制作用，输入Xi与输出Yi均为0或1，强调该神经元兴奋状态或抑制状态）**



用MP模型进行二元布尔逻辑运算：

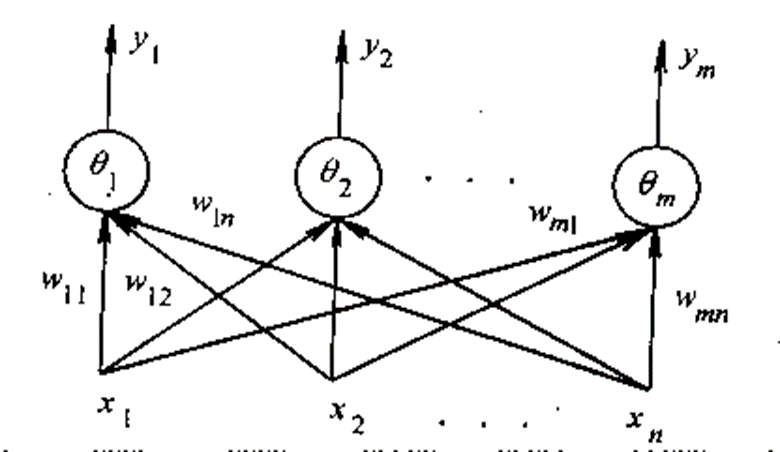


用MP模型进行三元布尔逻辑运算：



### Perceptron (Rosenblatt in 1957)

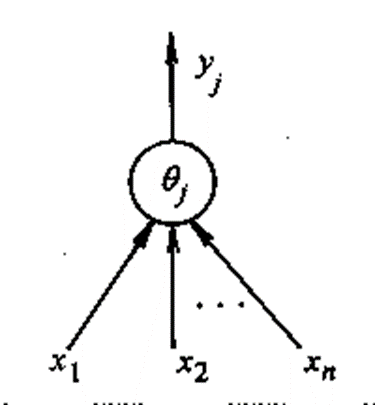
**单层感知器模型**

****

总输入；权值Wji：xi到yj的权值；输出：yj (j=1,2,…,m)输出为0或1。

结构基于MP模型，可以理解为多个MP模型构成的输入与输出之间的简单全连接。

单个神经元分析：

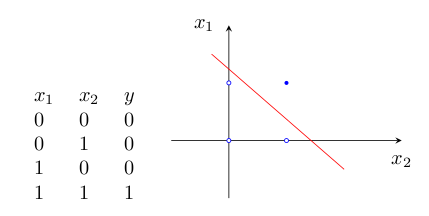
Wj2

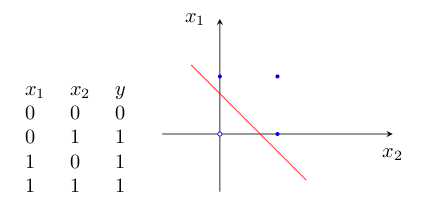
Wj1

Wjn

输入X=（X1,X2,…,Xn）可视为N维空间中的某个点。而感知器将点X分成了两类。

单层感知器如何解决线性问题：（与、或问题实质是线性分类问题）

逻辑与：

逻辑或：

对于感知器模型：

Yj=1, ωj1x1+ωj2x2−θj>0

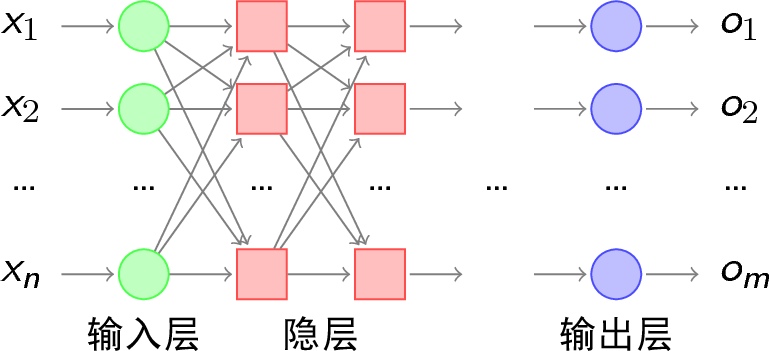
Yj=-1, ωj1x1+ωj2x2−θj<0

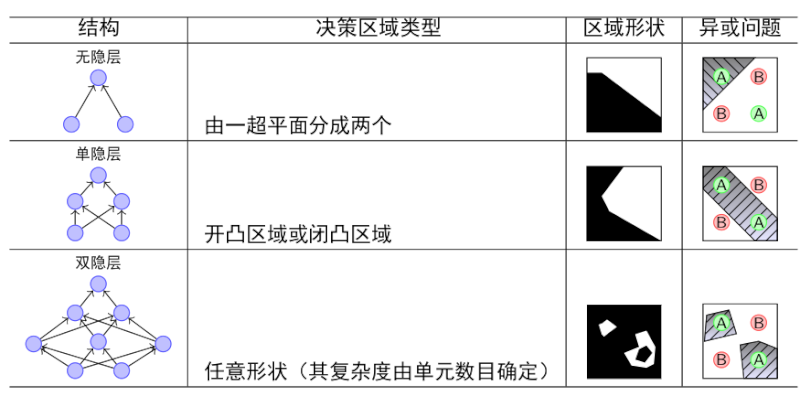
方程ωj1x1+ωj2x2−θj=0确定的直线就是二维输入样本空间上的一条分界线。将其权值进行学习调整，即可求出逻辑与和逻辑或问题中的直线（不唯一），这样，可以规定当给入输入后，若Yj=1，则此时点X（X1,X2）在直线上方，反之当Yj=-1，点X在直线下方，实现了对点X的线性分类。

Mnisky数学证明（穷举即可）了单层感知器模型无法解决异或问题（非线性分类问题），使人工神经网络发展陷入低潮，后来的多层感知器模型则可以解决这个问题。

**多层感知器模型**

结构如下：输出Oi等同Yi，图片换了个名字。



比较单层感知器和多层感知器的分类能力：（神经网络层数一般规定为隐含层数+1）。

由上图可以看出，随着隐层层数的增多，凸域可以被划分成任意的形状，因此可以解决任何复杂的分类问题。实际上，Kolmogorov理论指出：双隐层感知器就足以解决任何复杂的分类问题。基于不同的学习算法，衍生出不同网络。

Error Back Propagation算法简称BP算法，以BP算法实现的多层感知器网络就是BP神经网络。即BP神经网络本质=BP学习算法+多层感知器网络。

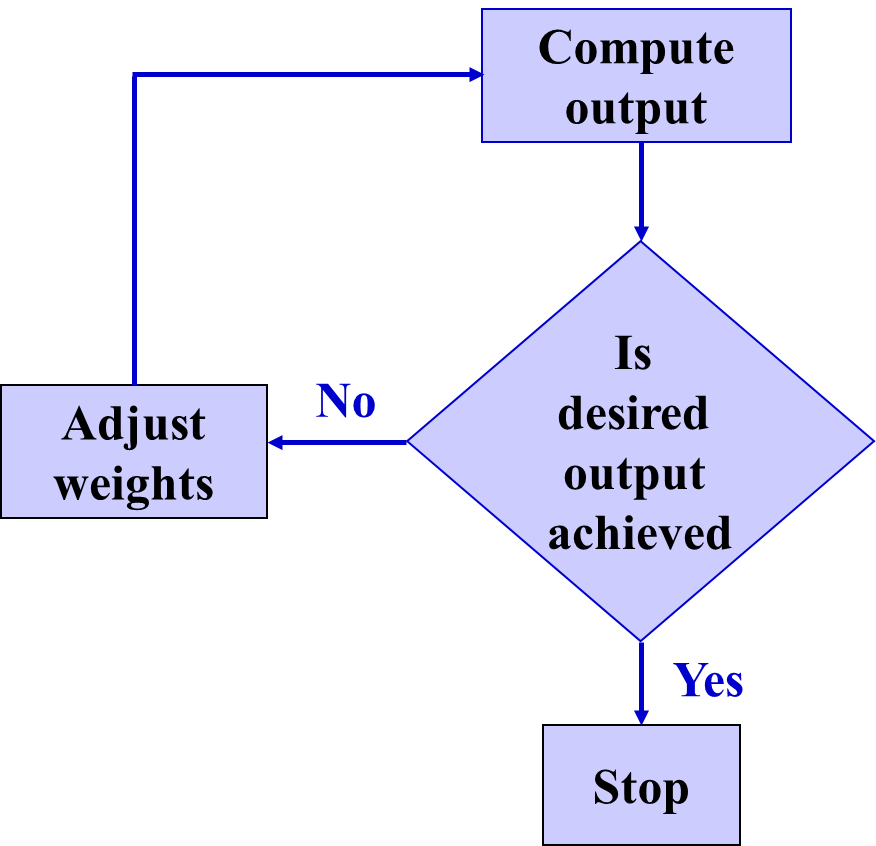
**MP模型与感知器模型的不同：**

MP模型不具备学习能力，没有学习算法调整权值。

感知器模型具备学习能力，配合相应的学习算法可以对权值进行调整，学习到的信息存储在权值上。感知器实际上是在MP模型的基础上加上学习功能，使其权值可以调节的产物。

### Learning rules

一般的训练（学习）过程：学习的过程，就是调整权值的过程。



学习类型分为有监督式（有导师式）的学习和无监督式（无导师式）的学习

**有监督式（有导师式）学习：**通过已有的训练样本（即已知数据以及其对应的输出）来训练，从而得到一个最优模型，再利用这个模型将所有新的数据样本映射为相应的输出结果，对输出结果进行简单的判断从而实现分类的目的，那么这个最优模型也就具有了对未知数据进行分类的能力。

**无监督式（无导师式）学习：**事先没有任何训练数据样本，需要直接对数据进行建模。样本数据类别未知，需要根据样本间的相似性对样本集进行分类（聚类，clustering）试图使类内差距最小化，类间差距最大化。通俗点将就是实际应用中，不少情况下无法预先知道样本的标签，也就是说没有训练样本对应的类别，因而只能从原先没有样本标签的样本集开始学习分类器设计。

**两者的不同点：**

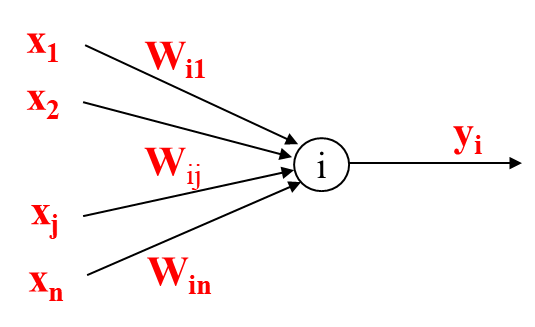
1.有监督学习方法必须要有训练集与测试样本。在训练集中找规律，而对测试样本使用这种规律。而非监督学习没有训练集，只有一组数据，在该组数据集内寻找规律。

2.有监督学习的方法就是识别事物，识别的结果表现在给待识别数据加上了标签。因此训练样本集必须由带标签的样本组成。而非监督学习方法只有要分析的数据集的本身，预先没有什么标签。如果发现数据集呈现某种聚集性，则可按自然的聚集性分类，但不予以某种预先分类标签对上号为目的。

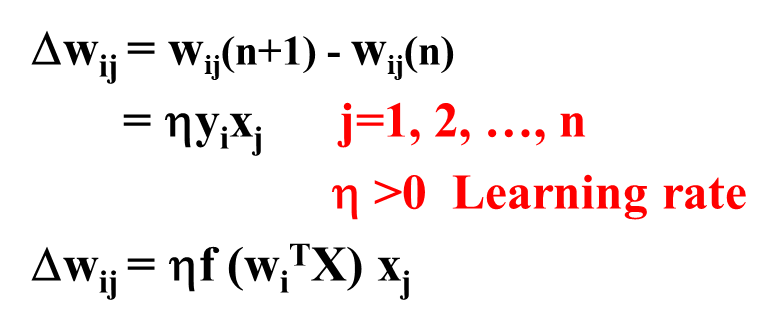
3.非监督学习方法在寻找数据集中的规律性，这种规律性并不一定要达到划分数据集的目的，也就是说不一定要“分类”。这一点是比有监督学习方法的用途要广。譬如分析一堆数据的主分量，或分析数据集有什么特点都可以归于非监督学习方法的范畴。

学习算法：如何更新权值即求**Δwij**的算法

**1.Hebb Learning Rule (无监督式学习)**

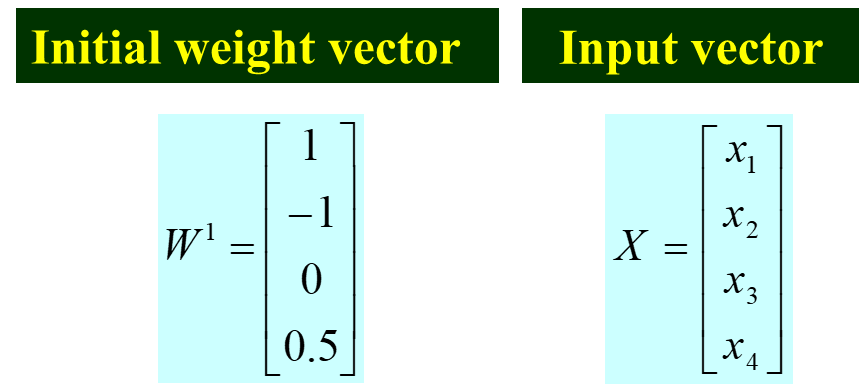


对于上图第i个神经元，其权值的变化量**Δwij**的计算方法如下:



举例：

1.初始化权值。



2.给定输入：

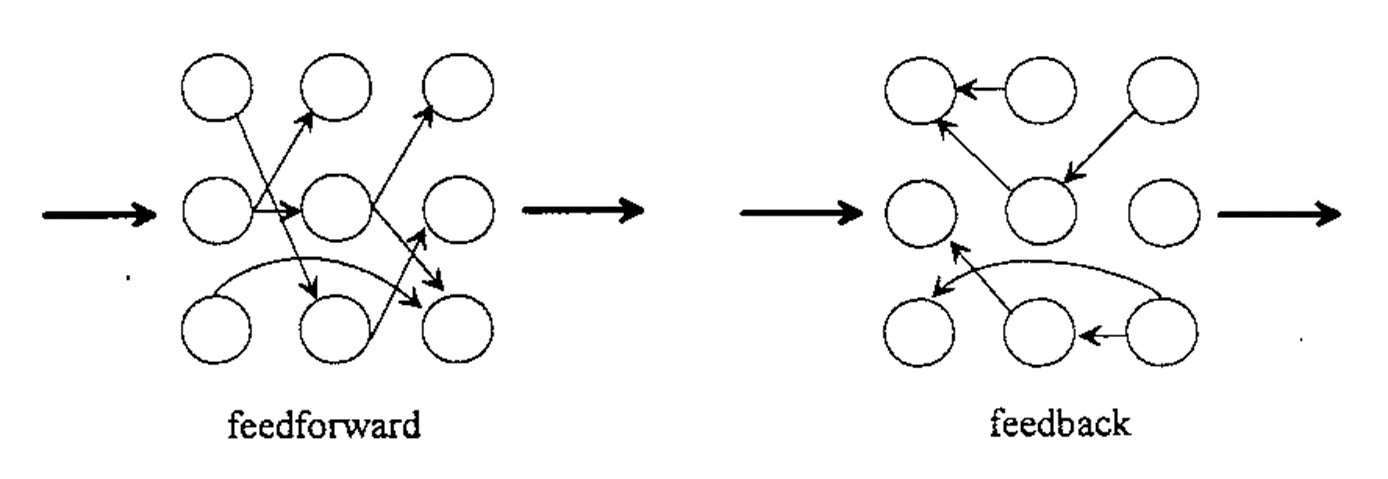


3.设定超参数：



4.迭代计算：

### 前馈神经网络与反馈神经网络



**前馈神经网络：**简称前馈网络，是人工神经网络的一种。在此种神经网络中，各神经元从输入层开始，接收前一级输入，并输出到下一级，直至输出层。整个网络中无反馈，可用一个有向无环图表示。采用一种单向多层结构。其中每一层包含若干个神经元，同一层的神经元之间没有互相连接，层间信息的传送只沿一个方向进行。其中第一层称为输入层。最后一层为输出层．中间为隐含层，简称隐层。隐层可以是一层。也可以是多层。代表网络有感知器网络，BP神经网络，RBF网络等。

**反馈神经网络：**反馈型神经网络是一种从输出到输入具有反馈连接的神经网络，其结构比前馈网络要复杂得多。每个神经元同时将自身的输出信号作为输入信号反馈给其他神经元，需要工作一段时间才能达到稳定。典型的反馈型神经网络有：Elman网络和Hopfield网络。

**两者区别：**

前馈型神经网络取连续或离散变量，一般不考虑输出与输入在时间上的滞后效应，只表达输出与输入的映射关系；反馈型神经网络可以用离散变量也可以用连续取值，考虑输出与输入之间在时间上的延迟，需要用动态方程来描述系统的模型。

前馈型神经网络的学习主要采用误差修正法（如BP算法），计算过程一般比较慢，收敛速度也比较慢；而反馈型神经网络主要采用Hebb学习规则，一般情况下计算的收敛速度很快。

反馈网络也有类似于前馈网络的应用，并且在联想记忆和优化计算方面的应用更显特点。

### 1962 Adaline (Adaptive linear element) (Widrow)

### 1969 Book “Perceptron” (Minsky and Papert)

### 1982 Hopfield network