人工神经网络总结（ANN）

人工神经网络是一个具有学习能力的系统，可以发展知识，以至超过设计者原有的知识水平。通常，它的学习(或训练)方式可分为两种，一种是有监督(supervised)学习，另一种是无监督(unsupervised)学习。

在人工神经网络设计及应用研究中，通常需要考虑三个方面的内容，即神经元激活函数、神经元之间的连接形式和网络的学习(训练)。

神经网络的学习策略

(1)、误差修正型规则：是一种有监督的学习方法，根据实际输出和期望输出的误差进行网络连接权值的修正，最终网络误差小于目标函数达到预期结果。

误差修正法，权值的调整与网络的输出误差有关， 它包括δ学习规则、Widrow-Hoff学习规则、感知器学习规则和误差反向传播的BP(Back Propagation)学习规则等。





(2)、竞争型规则：无监督学习过程，网络仅根据提供的一些学习样本进行自组织学习，没有期望输出，通过神经元相互竞争对外界刺激模式响应的权利进行网络权值的调整来适应输入的样本数据。

对于无监督学习的情况，事先不给定标准样本，直接将网络置于“环境”之中，学习(训练)阶段与应用(工作)阶段成为一体。

(3)、Hebb型规则：利用神经元之间的活化值(激活值)来反映它们之间联接性的变化，即根据相互连接的神经元之间的活化值(激活值)来修正其权值。

在Hebb学习规则中，学习信号简单地等于神经元的输出。Hebb学习规则代表一种纯前馈﹑无导师学习。该学习规则至今在各种神经网络模型中起着重要作用。典型的应用如利用Hebb规则训练线性联想器的权矩阵。

(4)、随机型规则：在学习过程中结合了随机、概率论和能量函数的思想，根据目标函数（即网络输出均方差）的变化调整网络的参数，最终使网络目标函数达到收敛值。

**激活函数**

如果没有激活函数呢，所有的神经网络只能解决线性问题，正是因为激活函数的存在，才让神经网络的能力扩展到了非线性问题。而激活函数的选择对网络的收敛速度有较大的影响，针对不同的实际问题，激活函数的选择也应不同。

其实有激活函数，我们还可以这样去想一想。我们模仿的是人的神经元，神经元在传递信号的时候，需要刺激到达一个阈值！这像不像我们的激活函数做到的一个效果呢？

比如说sigmoid函数、relu函数。

(1)、阈值函数：该函数通常也称为阶跃函数。当激活函数采用阶跃函数时，人工神经元模型即为MP模型。此时神经元的输出取１或０，反应了神经元的兴奋或抑制。

(2)、线性函数：该函数可以在输出结果为任意值时作为输出神经元的激活函数，但是当网络复杂时，线性激活函数大大降低网络的收敛性，故一般较少采用。

(3)、对数S形函数：对数S形函数的输出介于0~1之间，常被要求为输出在０～１范围的信号选用。它是神经元中使用最为广泛的激活函数。

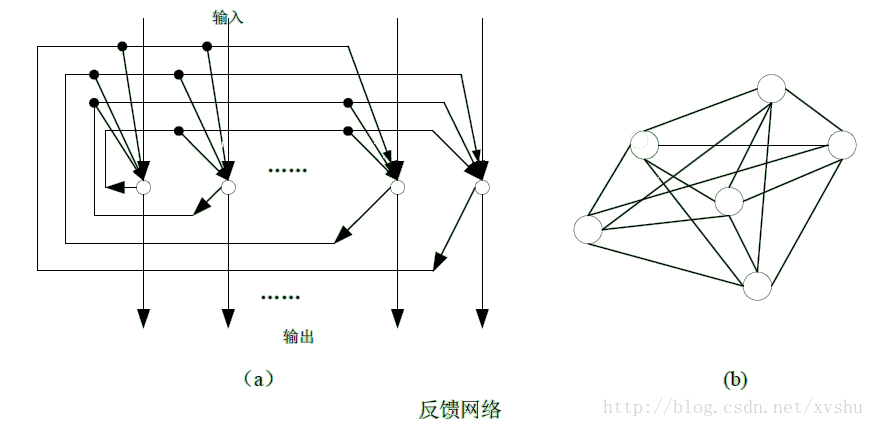
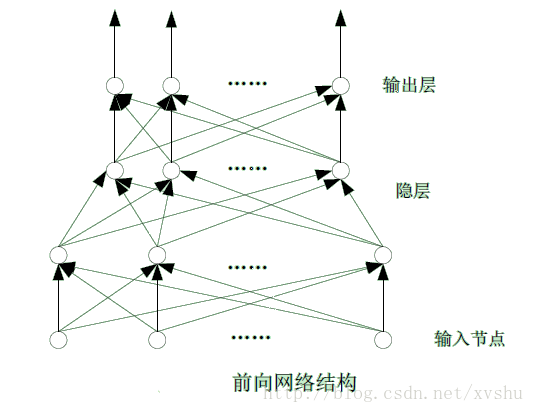
(4)、双曲正切S形函数：双曲正切S形函数类似于被平滑的阶跃函数，形状与对数S形函数相同，以原点对称，其输出介于-１~１之间，常常被要求为输出在-１~１范围的信号选用。

**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影响。互连模式种类繁多。

(1)、前馈网络：网络可以分为若干“层”，各层按信号传输先后顺序依次排列，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。前馈型网络可用一有向无环路图表示，如下图所示：

(2)、反馈网络：每个节点都表示一个计算单元，同时接受外加输入和其它各节点的反馈输入，每个节点也都直接向外部输出。在某些反馈网络中，各神经元除接受外加输入与其它各节点反馈输入之外，还包括自身反馈。有时，反馈型神经网络也可表示为一张完全的无向图，如上图b。图中，每一个连接都是双向的。这里，第i个神经元对于第j个神经元的反馈与第j至i神经元反馈之突触权重相等，也即wij=wji。



**人工神经网络的分类：**

按性能分：连续型和离散型网络，或确定型和随机型网络。

按拓扑结构分：前向网络和反馈网络。

前馈网络有自适应线性神经网络(AdaptiveLinear，简称Adaline)、单层感知器、多层感知器、BP等。

反馈网络有Hopfield、Hamming、BAM等。

反馈网络，网络内神经元间有反馈，可以用一个无向的完备图表示。这种神经网络的信息处理是状态的变换，可以用动力学系统理论处理。系统的稳定性与联想记忆功能有密切关系。Hopfield网络、波耳兹曼机均属于这种类型。

按学习方法分：有教师(监督)的学习网络和无教师(监督)的学习网络。

按连接突触性质分：一阶线性关联网络和高阶非线性关联网络。

## 反向传播模型

在我看来单层感知机、多层感知机只是一个浅层网络的一个特例，或者也可以看作是反向传播模型的一个特例。



（单层感知机↑）





（bp↑）

反向传播网络的学习算法：B-P算法的学习目的是对网络的连接权值进行调整，使得调整后的网络对任一输入都能得到所期望的输出。

学习过程由正向传播和反向传播组成。

正向传播用于对前向网络进行计算，即对某一输入信息，经过网络计算后求出它的输出结果。

反向传播用于逐层传递误差，修改神经元间的连接权值，以使网络对输入信息经过计算后所得到的输出能达到期望的误差要求。

B-P算法的优缺点：

优点：

理论基础牢固，推导过程严谨，物理概念清晰，通用性好等。所以，它是目前用来训练前向多层网络较好的算法，有很多的算法都是基于bp的。

缺点：

(1)、该学习算法的收敛速度慢；

(2)、网络中隐节点个数的选取尚无理论上的指导； 当层数太多的时候又可能对于前面的几层参数的调整程度太小。

(3)、从数学角度看，B-P算法是一种梯度最速下降法，这就可能出现局部极小的问题。当出现局部极小时，从表面上看，误差符合要求，但这时所得到的解并不一定是问题的真正解。所以B-P算法是不完备的。

改进算法：

(1)、增加动量项：标准BP算法在调整权值时，只按t时刻误差的梯度下降方向调整，而没有考虑t时刻以前的梯度方向，从而常使训练过程发生振荡，收敛缓慢。为了提高训练速度，可以在权值调整公式中加一动量项。大多数BP算法中都增加了动量项，以至于有动量项的BP算法成为一种新的标准算法。

(2)、可变学习速度的反向传播算法（variable learning rate back propagation，VLBP）：多层网络的误差曲面不是二次函数。曲面的形状随参数空间区域的不同而不同。可以在学习过程中通过调整学习速度来提高收敛速度。技巧是决定何时改变学习速度和怎样改变学习速度。可变学习速度的VLBP算法有许多不同的方法来改变学习速度。

(3)、学习速率的自适应调节：可变学习速度VLBP算法，需要设置多个参数，算法的性能对这些参数的改变往往十分敏感，另外，处理起来也较麻烦。此处给出一简洁的学习速率的自适应调节算法。学习率的调整只与网络总误差有关。学习速率η也称步长，在标准BP中是一常数，但在实际计算中，很难给定出一个从始至终都很合适的最佳学习速率。从误差曲面可以看出，在平坦区内η太小会使训练次数增加，这时候希望η值大一些；而在误差变化剧烈的区域，η太大会因调整过量而跨过较窄的“凹坑”处，使训练出现振荡，反而使迭代次数增加。为了加速收敛过程，最好是能自适应调整学习率η，使其该大则大，该小则小。比如可以根据网络总误差来调整.

(4)、引入陡度因子—-防止饱和：误差曲面上存在着平坦区。其权值调整缓慢的原因在于S转移函数具有饱和特性造成的。如果在调整进入平坦区后，设法压缩神经元的净输入，使其输出退出转移函数的饱和区，就可改变误差函数的形状，从而使调整脱离平坦区。实现这一思路的具体作法是在转移函数中引进一个陡度因子。

**而在bp之上，我们加几层rbm就可以得到一种新的方法，DBN深度信念网络。**

**Rbm是通过引入能量函数进行一个训练，**主要是构建了(v,h)之间的联合分布，根据这个联合分布我们可以用来做分类、降维、参数预训练等等。

Dbn通过几层rbm进行一个预训练，之后再通过一个bp算法对于参数进行微调。

传说autoEncoder和RBM预训练初始化参数会很好地提升模型的结果

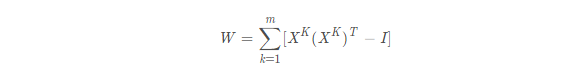
## Hopfield模型：

网络系统具有若干个稳定状态，对应了能量函数的局部最小值，当网络从某一初始状态开始运动，网络系统总可以收敛到某一个稳定的平衡状态；

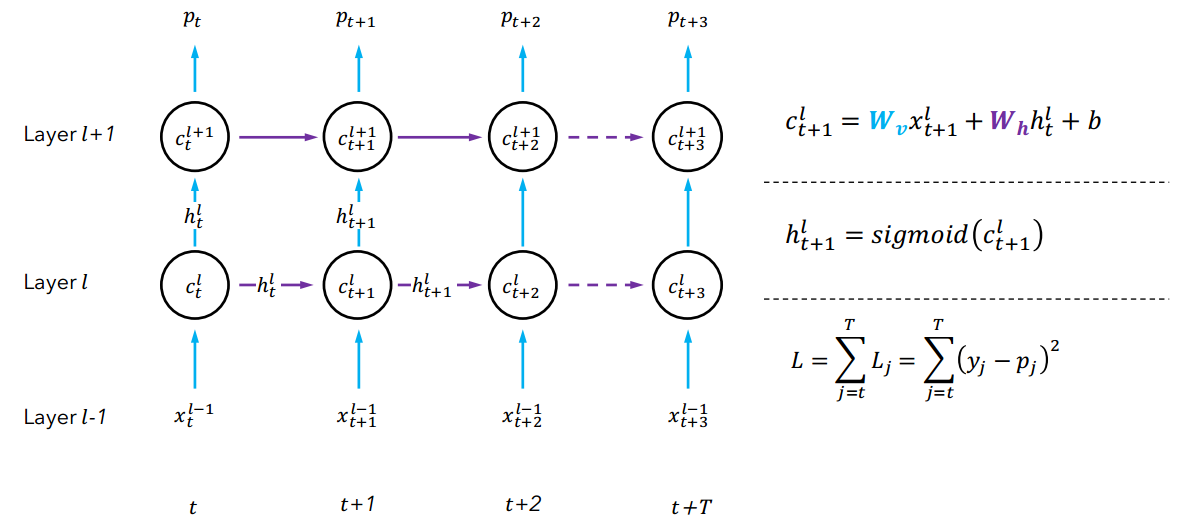
系统稳定的平衡状态可以通过设计网络的权值而被存储到网络中；

我们可以根据能量极小点设计满足条件的网络参数，方法就是设计一组方程，每一个方程表示当其中一个输出神经元变化的时候，能量变化量的大小，令每一个方程都大于等于，即不管任一个神经元如何变化，总的能量都不会变小，根据一系列方程组就可以求得我们想要的参数。

Hopefield网络可以用于联想记忆和计算，当给定权重矩阵 W，求网络最终的稳定状态时，其用于计算；反过来则是用于联想记忆。

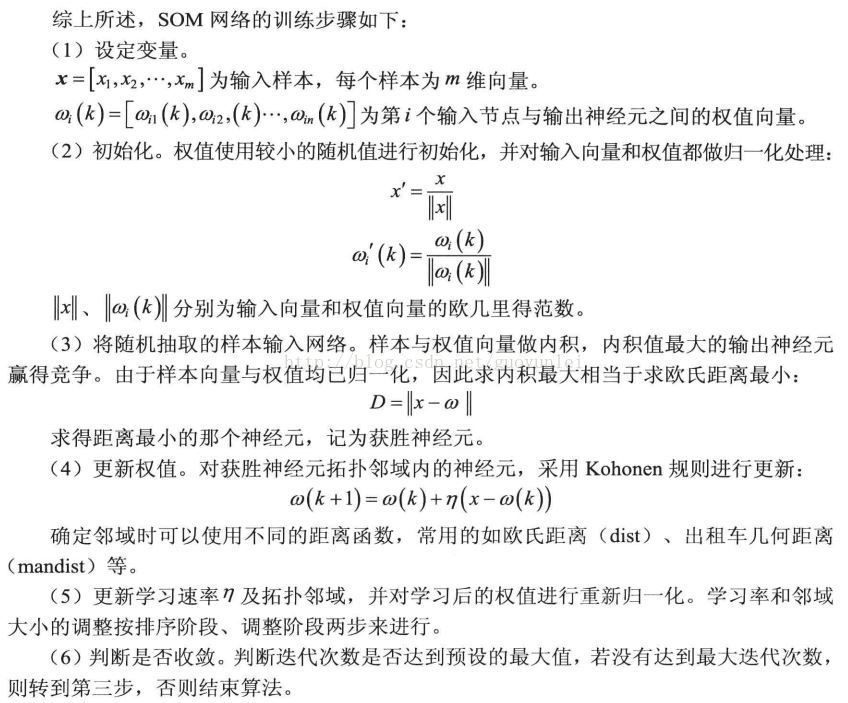
**Bptt**



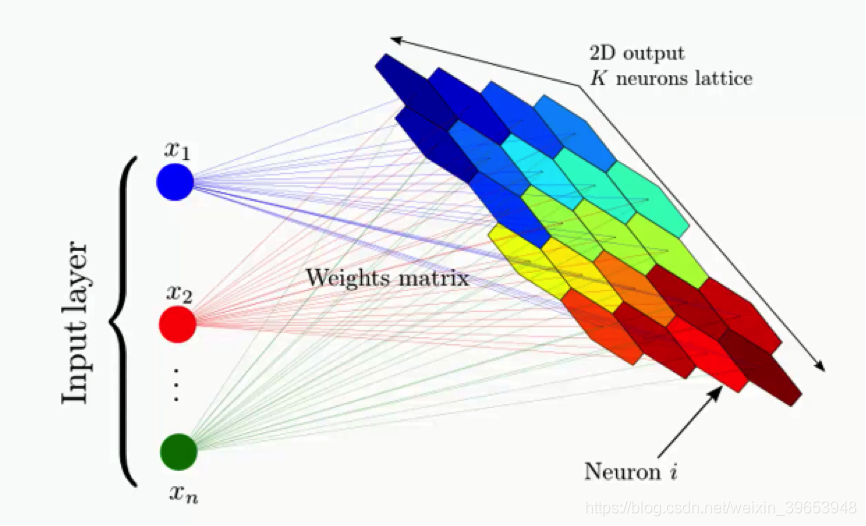
Bptt是一个带有自环的、反馈型的，也可以把他看作是按照时间展开的。

所以在我看来呢是一个对于bp 的改善。

**自组织特征映射网络（SOFM）**



自组织映射（Self-Organizing Map ，SOM）或自组织特征映射（Self-Organizing Feature Map ，SOFM） 是一种类型的人工神经网络（ANN），其使用已训练的无监督学习以产生的一个低维（通常为二维），离散表示训练样本的输入空间（称为Map），因此是一种降维方法。下图示例显示了如何降低复杂数据集的维度。



1. 自组织特征映射网络

自组织特征映射网络(Self-Organizing Feature Map)只有两层：输入层和竞争层，竞争层神经元的排列有多种形式：一维线阵、二维平面、三维栅格等等。

SOFM网的运行分训练和工作两个阶段。在训练阶段，对网络随机输入训练集中的样本，对某个特定的输入模式，输出层会有某个神经元产生最大响应而获胜，而在训练开始阶段，输出层哪个位置的神经元将对那类输入模式产生最大响应是不确定的。当输入模式的类别改变时，二维平面的获胜神经元也会改变。获胜神经元周围的神经元因侧向相互兴奋作用产生较大响应，于是获胜神经元以及其优胜邻域内的所有神经元所连接的权向量均向输入向量的方向作程度不同的调整，调整力度依邻域内各神经元距获胜神经元的远近而逐渐衰减。网络通过自组织方式，用大量训练样本调整网络的权值，最后使输出层各神经元成为特定模式类敏感的神经细胞，对应的内星权向量成为各输入模式类的中心向量。并且当两个模式类的特征接近时，代表这两类的神经元在位置也接近。从而在输出层形成能够反映样本模式类分布情况的有序特征图。

SOFM网训练结束后，输出层各神经元与各输入模式类的特定关系就完全确定了，因此可用作模式分类器。当输入一个模式时，网络输出层代表该模式类的特定神经元将产生最大响应，从而将该输入自动归类。应当指出的是，当向网络输入的模式不属于网络训练时见过的任何模式类时，SOFM网只能将它归入最接近的模式类。

