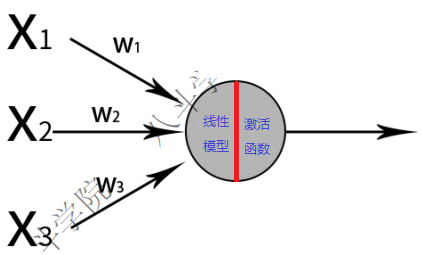
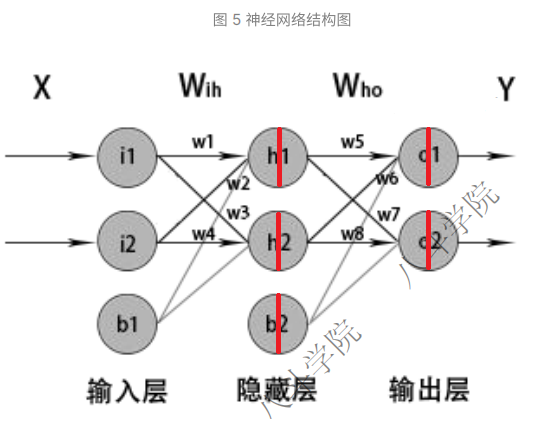
# NN神经网络原理

## NN的结构



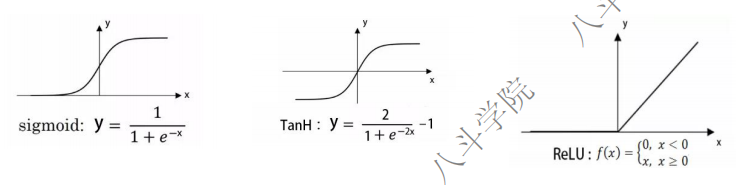
我们⽬前使⽤的神经元内部的运算通常有两个部分组成，第⼀部分是所谓的“线性模型”，另外⼀部分是“激活函数”，其作用往往是将线性结果转换为符合特定需求的非线性结果.



* 隐藏层: 这⼀层可以有多层，在⽐较深的⽹络中，隐藏层能达到数⼗上百层
* 输出层: 是最后⼀层，⽤来输出整个⽹络计算后的结果，这⼀层可能是⽐较复杂的类型的值或者向量，根据不同的需求输出层的构造也是不同的。
* 前⼀层的神经元节点连接了下⼀层的所有节点,这种前后层相互连接的⽹络也叫作全连接神经⽹络，
* 由于数据是⼀层⼀层从输⼊⾄输出传播的，也叫作前馈神经⽹络。
* 前向传播(forward)
* 反向传播(backward)
* 房价预测问题,输出层应该有一个神经元, 该神经元的输出对应房价的预测值
* 分类问题,输出层应该有二个或多个神经元,每个神经元的输出对应一个类别的预测概率

## **NN常用激活函数**

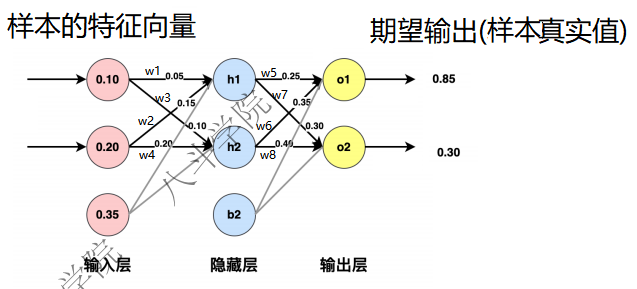
神经网络中常用到的三种激活函数:



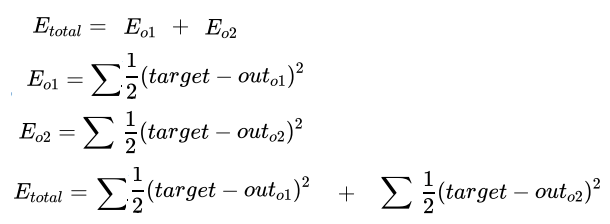
* sigmoid 函数: 是将线性模型的计算结果投射到 0 到 1 之间, 常用与二分类问题,多分类问题需要⽤到 softmax 函数
* TanH 函数: 是将线性模型的计算结果投射到-1 到 1 之间
* ReLU函数: 是将线性模型计算结果⼩于 0 的部分投射为 0，⼤于等于 0 的部分投射为计算结果本⾝。

# NN损失函数 & 梯度

## MSE均方误差损失



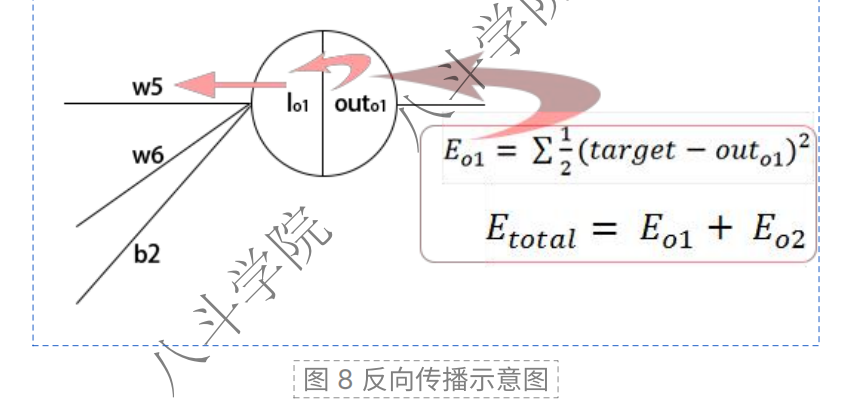
针对上面的这个简单神经网络, 如果用MSE均方误差衡量真实值与预测值之间的差异, 损失函数计算如下:





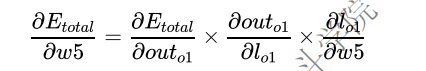
### 梯度计算

在模型训练 (参数估计) 期间, 计算损失函数的梯度目的是用于更新模型参数, 以减少损失函数的值,

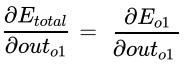


对于权重参数W5来说，它的变化对损失函数的影响, 即损失函数对w5的偏导数等于:

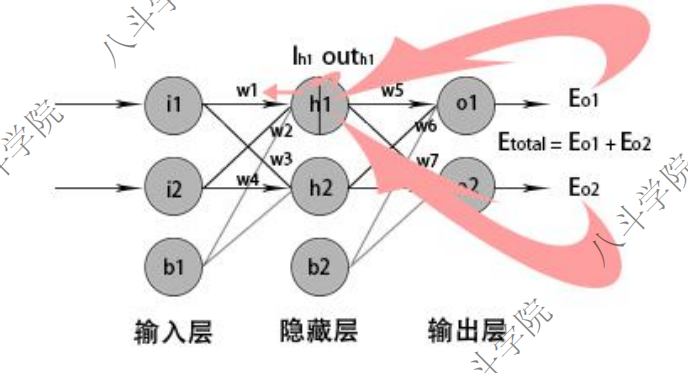


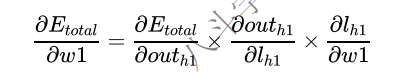


因为out-o1只对Eo1有影响, 所以:

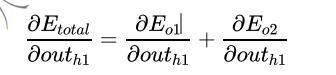


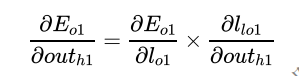
对于权重参数W1来说，它的变化对损失函数的影响, 即损失函数对w1的偏导数等于:

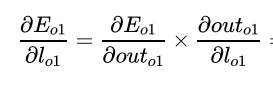




与上面w5不同的是, 这里out-h1对Eo1和Eo2都有影响. 所以:







## 分类交叉熵损失

