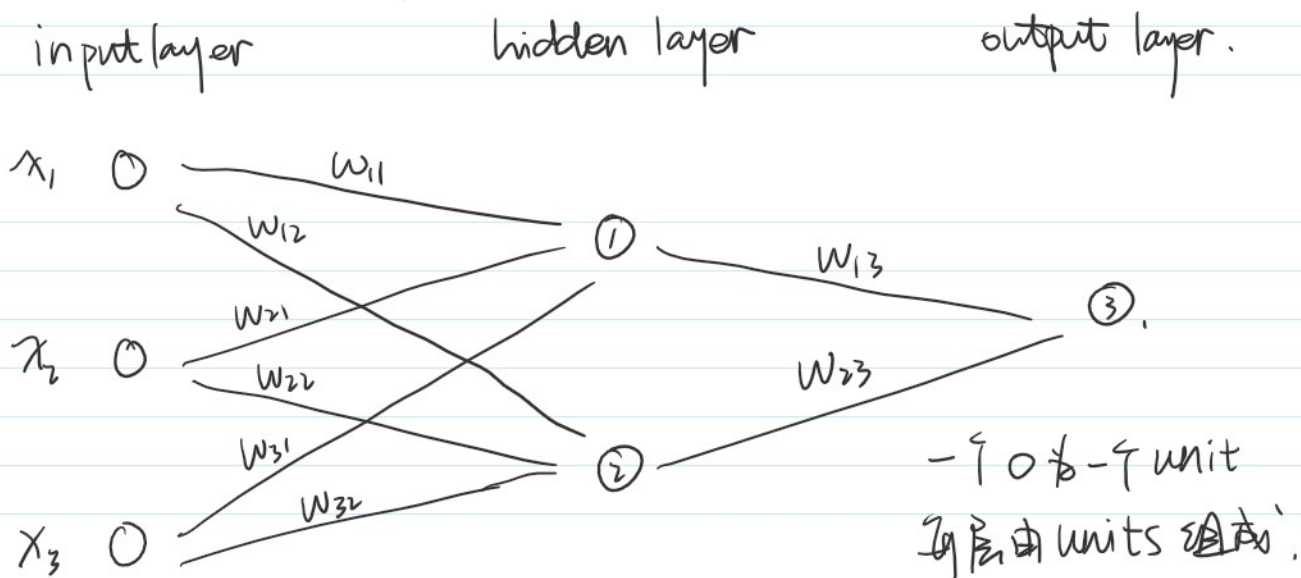


## 神经网络.



输入层输入的是训练集的特征向量.

特征向量在输入之前通常先进行标准化(normalize)  $xy \sim 1$  之间.

神经网络可以解决分类问题也可以解决回归问题.  
classification                  regression

输入层的单元设置.

特征向量  
如特征值  $A = (a_0, a_1, a_2)$  可用3个输出单元表示  $A$ .

若  $A = a_0$  则  $A = (1, 0, 0)$

若  $A = a_2$  则  $A = (0, 0, 1)$

以此类推.

输出层的单元设置.

对于分类问题.

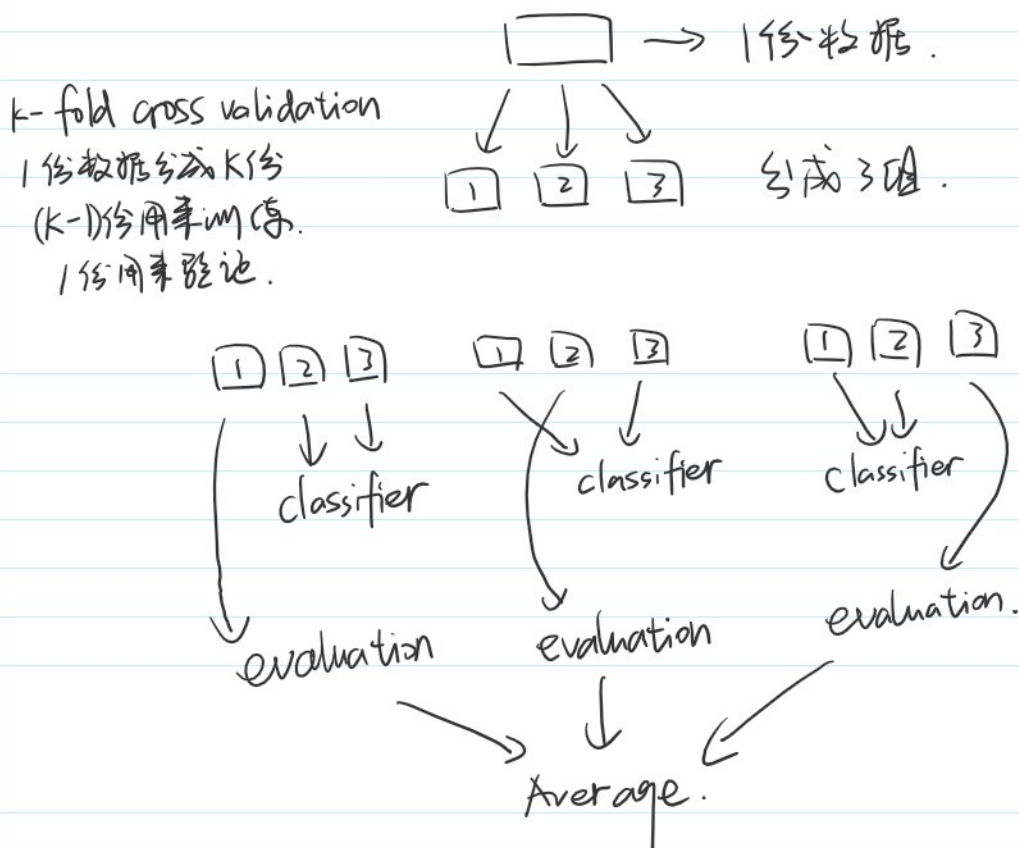
若为2类, 则用一个输出单元表示.

如黑白, 则黑为1, 白为0.

如黑白, 黑为1, 白为0.

若多于2类, 则每一类用一个输出单元表示.

## 交叉验证法 cross-validation.



## Backpropagation 算法.

(正向推算, 反向更新).

输入: 数据集, learning rate, 一个多层前馈的神经网络.

初始化 weight 及 bias

(随机初始化为  $-1 \sim 1$  或  $-0.5 \sim 0.5$  之间).

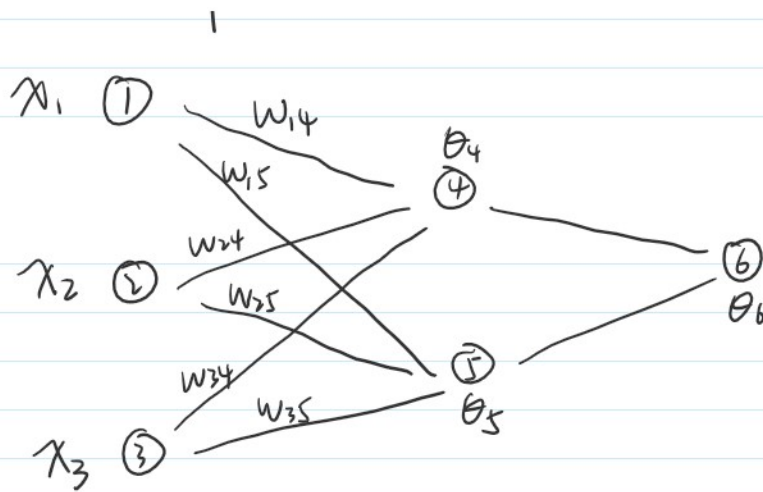
### 正向推进:

$$I_j = \sum_{i=1}^n w_{ij} O_i + \theta_j$$

(每一个单元的值, 等于上一层所有单元的值乘以相应的权重, 是bott bias)

### 非线性替代





$x_1$	$x_2$	$x_3$	$x_6$	$\theta_4$	$\theta_5$	$\theta_6$
1	0	1	1	-0.4	0.2	0.1

$w_{14}$	$w_{15}$	$w_{24}$	$w_{25}$	$w_{34}$	$w_{35}$	$w_{46}$	$w_{56}$
0.2	-0.3	0.4	0.1	-0.5	0.2	-0.3	-0.2

正向传播.

$$I_4 = \overset{w_{14}}{(0.2)} \times \overset{x_1}{1} + \overset{w_{24}}{(0.4)} \times \overset{x_2}{0} + \overset{w_{34}}{(-0.5)} \times \overset{x_3}{1} - \overset{\theta_4}{0.4} = -0.1$$

$$O_4 = 1 / (1 + e^{-(-0.1)}) = 0.332 \quad // \text{非线性转换.}$$

$$I_5 = (-0.3) \times 1 + (0.1) \times 0 + (0.2) \times 1 + 0.2 = 0.1$$

$$O_5 = 1 / (1 + e^{-(0.1)}) = 0.525.$$

$$I_6 = \overset{w_{46}}{(-0.3)} \times \overset{O_4}{0.332} + \overset{w_{56}}{(-0.2)} \times \overset{O_5}{0.525} + \overset{\theta_6}{0.1} = -0.105$$

$$O_6 = 1 / (1 + e^{-(-0.105)}) = 0.474.$$

反向传播.

$$Err_6 = \overset{O_j}{0.474} \cdot (1 - \overset{O_j}{0.474}) \cdot (\overset{T_j - O_j}{-0.474}) = 0.1311$$

上一层的误差 × 权重并求和.

$$Err_5 = \overset{O_j}{0.525} \cdot (1 - \overset{O_j}{0.525}) \cdot \sum_k Err_k \cdot \overset{w_{jk}}{(-0.2)} = -0.0065$$

$$\text{Err}_4 = 0.332 \times (1 - 0.332) \times (0.1311) \times (-0.3) = -0.0087$$

$$\text{设置 } (l) = 0.9$$

$$W_{46} = \overset{W_{46}}{-0.3} + (l) \times \overset{\text{Err}_6}{(0.1311)} \times \overset{O_4}{(0.332)} = -0.261$$

$$W_{ij} = W_{ij} + (l) \cdot \text{Err}_j \cdot O_i$$

$$W_{56} = -0.2 + (0.9) \times (0.1311) \times (0.525) = -0.138$$

$$W_{14} = 0.2 + 0.9 \times (-0.0087) \times 1 = 0.192$$

$$W_{15} = -0.3 + 0.9 \times (-0.0065) \times 1 = -0.306$$

$$W_{x4} = 0.4 + 0.9 \times (-0.0087) \times 0 = 0.4$$

$$W_{x5} = 0.1 + 0.9 \times (-0.0065) \times 0 = 0.1$$

$$W_{34} = -0.5 + 0.9 \times (-0.0087) \times 1 = -0.508$$

$$W_{35} = 0.2 + 0.9 \times (-0.0065) \times 1 = 0.194$$

$$\theta_6 = \overset{\theta_6}{0.1} + (l) \times \overset{\text{Err}_j}{0.1311} = 0.218$$

$$\theta_5 = 0.2 + 0.9 \times (-0.0065) = 0.194$$

$$\theta_4 = -0.4 + 0.9 \times (-0.0087) = -0.408.$$