

James 在1890年描述了神经网络的基本原理： 大脑皮层每一点的活力是由其他点势能释放的综合效能产生的。这一势能同下面的因素有关： 相关其他点的兴奋次数；兴奋的强度； 与其他不相连的其他点所接受的能量。

人工神经元

人工神经元模拟生物神经元的行为对输入信号作权和运算

$$y = w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n$$

其中 x_1, x_2, \cdots, x_n 是输入, $w_0, w_1, w_2, \cdots, w_n$ 是权重, 而 y 是输出.

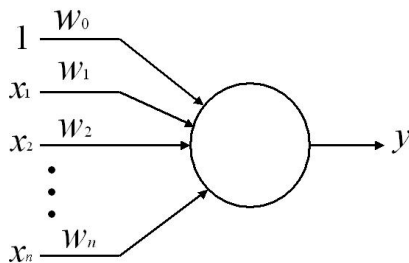


Figure: 一个神经元

Sigmoid 函数

定义一个没记忆的非线性函数 σ 作为激励函数来改变输出值

$$y = \sigma(w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n).$$

激励函数的选取依赖于应用问题.

Sigmoid 函数

定义一个没记忆的非线性函数 σ 作为激励函数来改变输出值

$$y = \sigma(w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n).$$

激励函数的选取依赖于应用问题. 如可用 sigmoid 函数

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

它的导数为

$$\sigma'(x) = \frac{e^{-x}}{(1 + e^{-x})^2}.$$

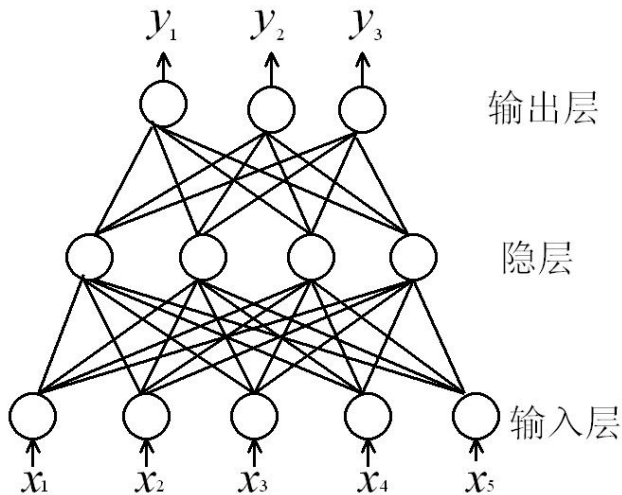


Figure: 前向型神经网络

我们考虑一个单隐层的 NN, 其中输入层有 n 个神经元, 输出层有 m 个神经元, 而隐含层有 p 个神经元. 那么在隐层神经元的输出为

$$x_i^1 = \sigma \left(\sum_{j=1}^n w_{ij}^0 x_j + w_{i0}^0 \right), \quad i = 1, 2, \dots, p.$$

因此输出层神经元的输出为

$$y_i = \sum_{j=1}^p w_{ij}^1 x_j^1 + w_{i0}^1, \quad i = 1, 2, \dots, m.$$

系数 $w_{ij}^0, i = 1, 2, \dots, p, j = 0, 1, \dots, n$ 和 $w_{ij}^1, i = 1, 2, \dots, m, j = 0, 1, \dots, p$ 称为网络权.