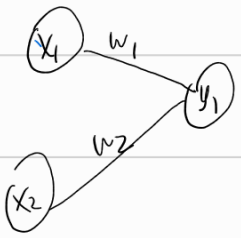
① 明确什么是线性和非线性，感知机是什么



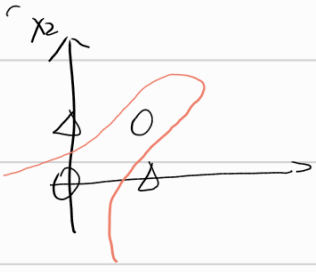
or门

x_1	x_2	y_1
0	0	0
0	1	1
1	0	1
1	1	1

↑ 线性可分(分割)

XOR门

x_1	x_2	y_1
0	0	0
0	1	1
1	0	1
1	1	0

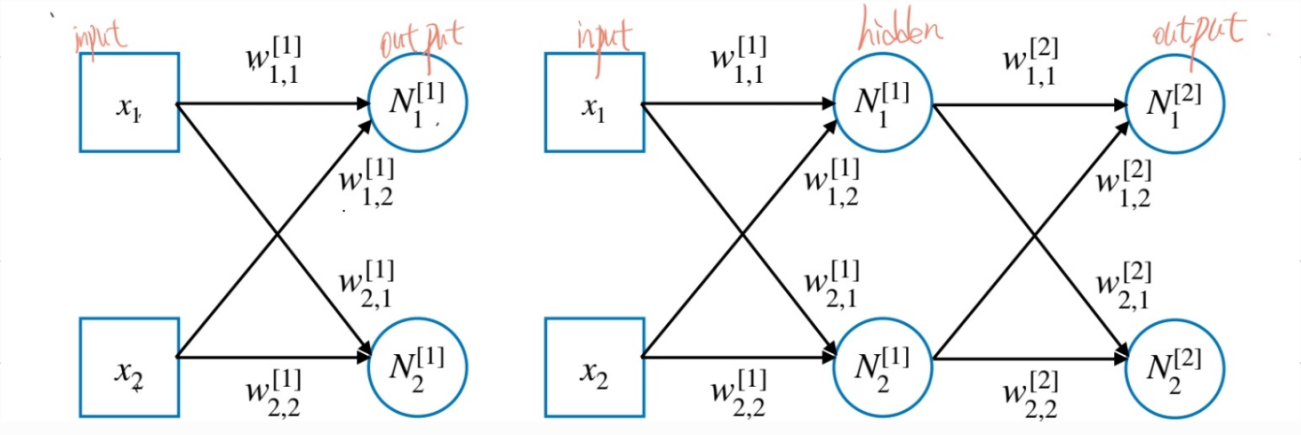


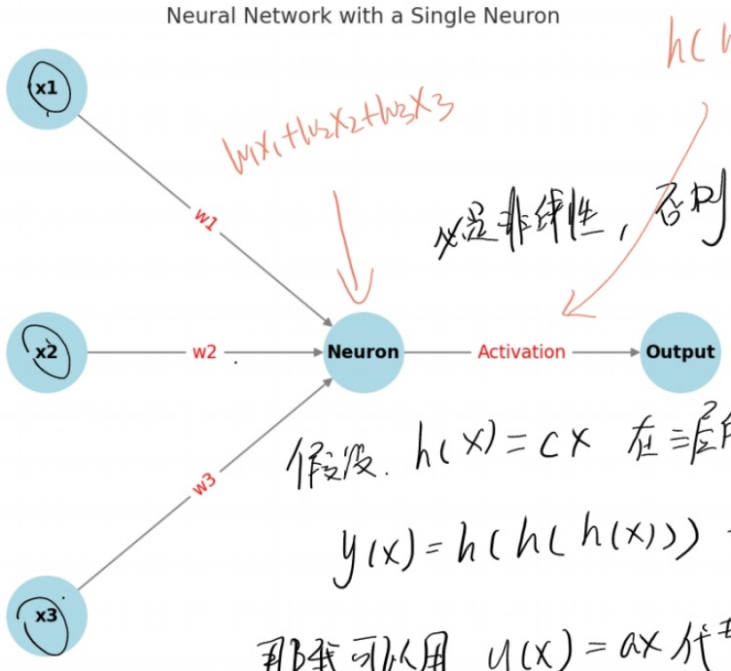
← 线性不可分(非线性)

⇓

单层的感知机做不了什么事情.

② 简单神经网络结构





$h(w_1x_1+w_2x_2+w_3x_3)=output$

$w_1x_1+w_2x_2+w_3x_3$

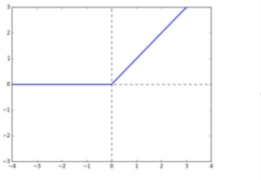
必须非线性，否则无论用一层代替多层

假设 $h(x)=cx$ 在三层的情况下

$y(x)=h(h(h(x)))=c\cdot c\cdot c\cdot x$

那我可以用 $y(x)=ax$ 代替 $a=c^3$

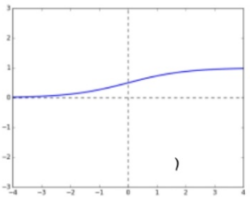
activation function



Rectified Linear Unit (ReLU)

$y=\max(0,z)$

$w_1x_1+w_2x_2$



Logistic sigmoid

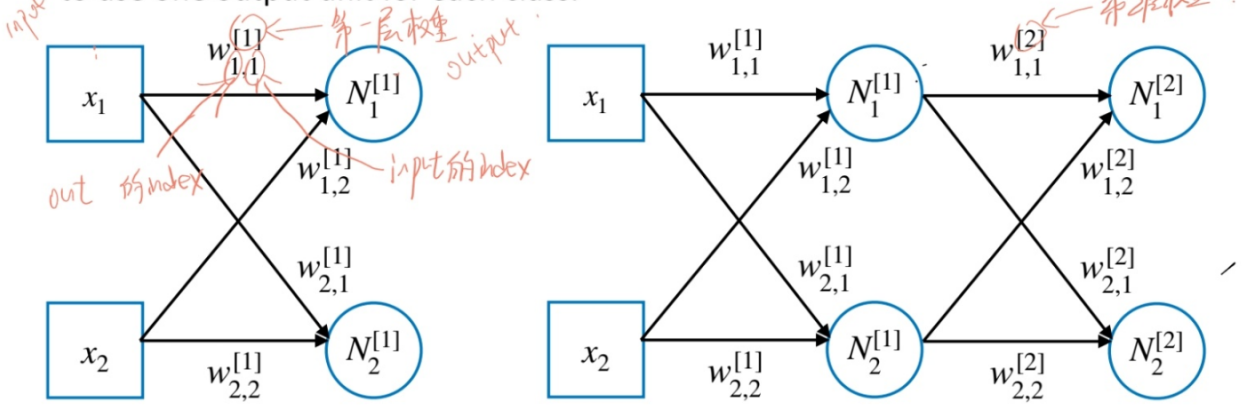
$$y=\frac{1}{1+e^{-z}}$$

这2种较常用

且后面的ppt会用到

认知 w的表示方式

- So far, we have considered only learning problems with a single output variable y, but **neural networks are often used in cases where multiple outputs are appropriate**.
  - For example, when learning to categorize images of handwritten digits—it is common to use one output unit for each class.

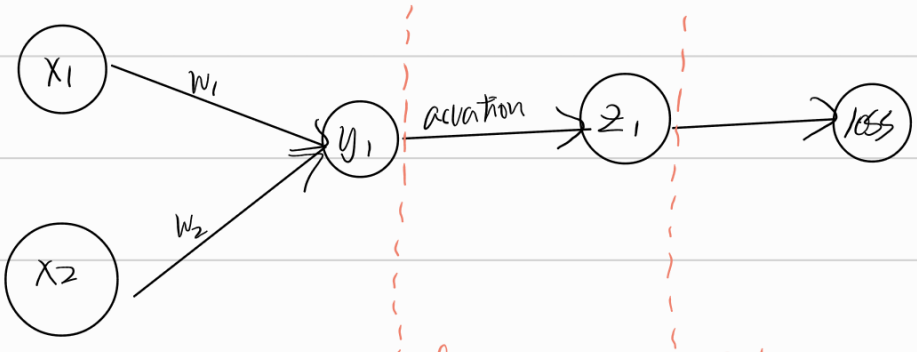


正向传播

① 从左到右计算每1个节点的值.

② 使用激活函数计算输入

③ 得到 z_1 值后与实际值比较，计算出 Loss



① $w_1x_1+w_2x_2=y_1$ ② $f(y_1)=z_1$ ③ 计算 $Loss(z_1)$