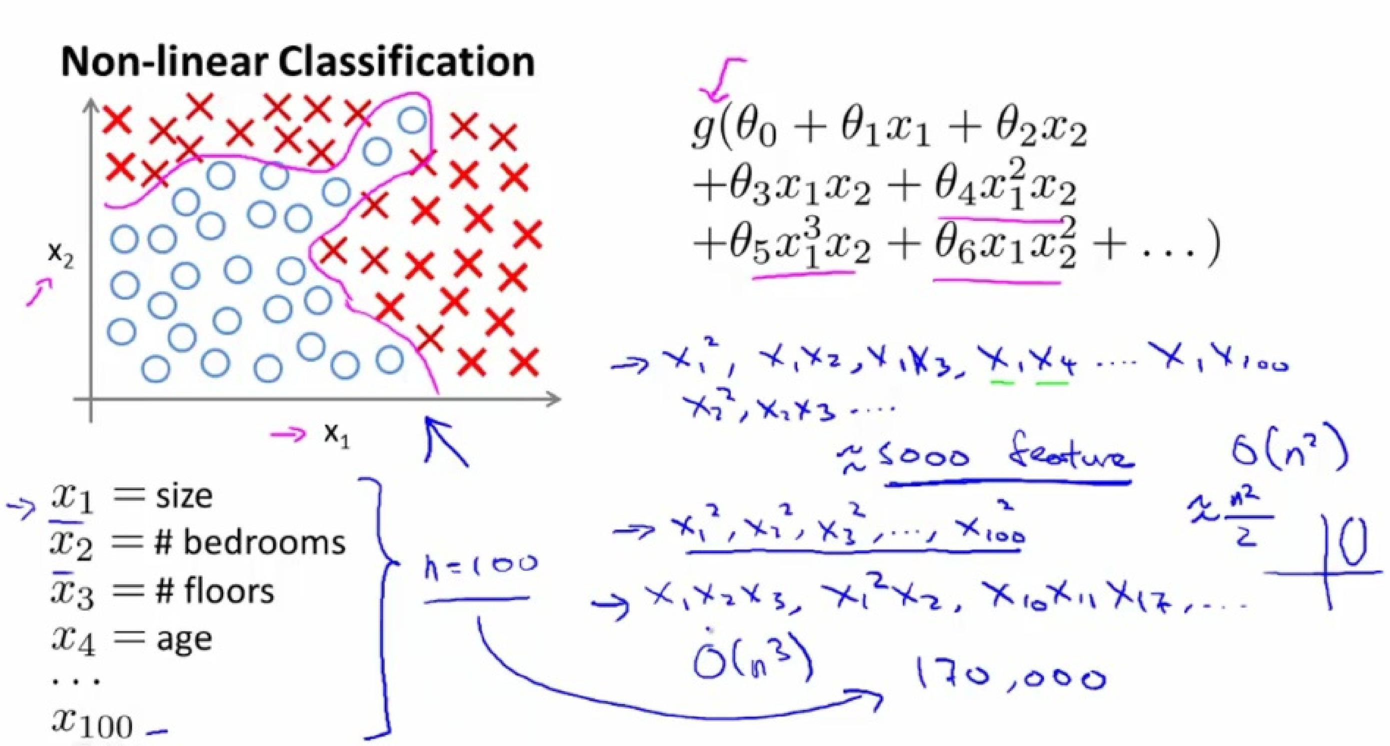
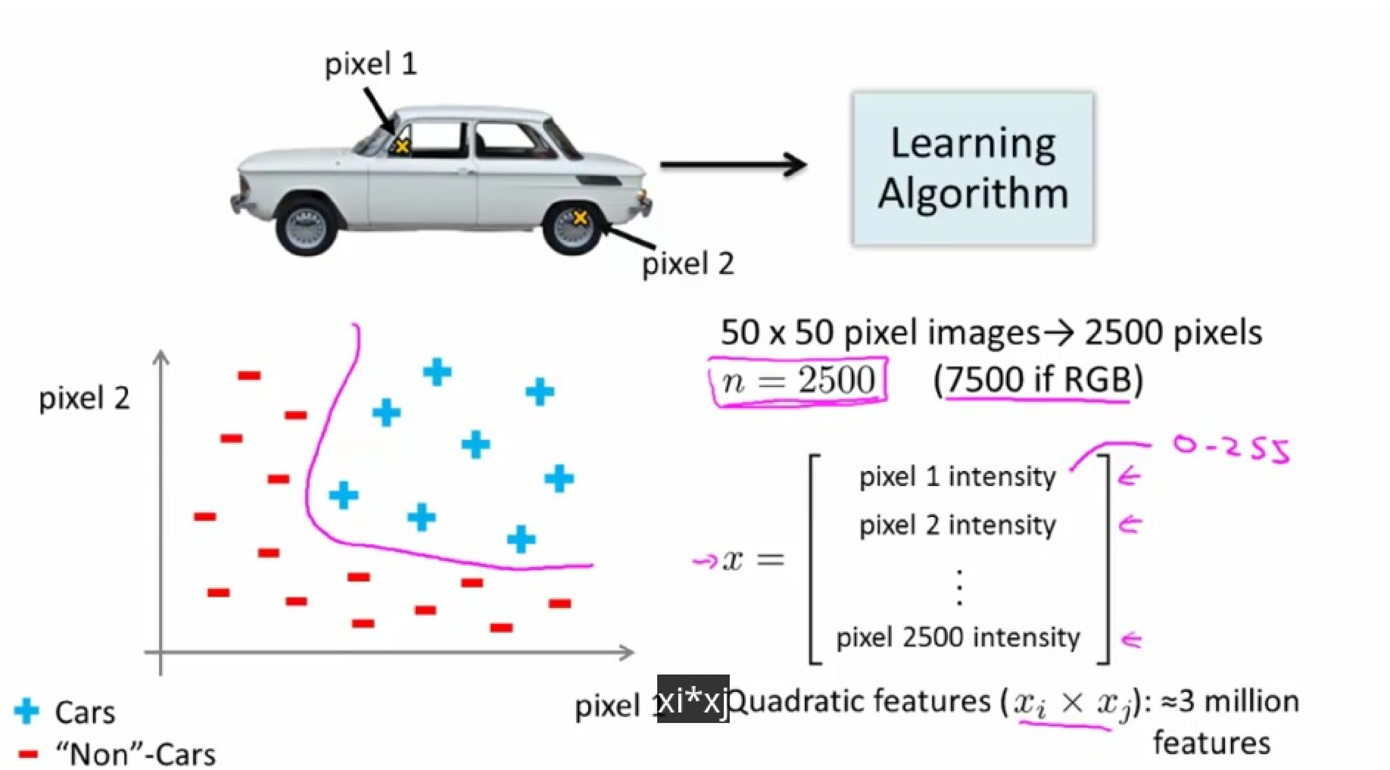
1. 当特征数量很多的时候，如果想包括其中所有的二次项是很难的。当然可以只包括其中部分的二次项，但是忽略太多的二次项会影响模型效果，另外还存在一个问题，除了二次项，三次项的数量更多。



检测汽车的学习算法特征数量：



因此简单的增加二次项或者三次项不是一个解决非线性回归问题的办法。

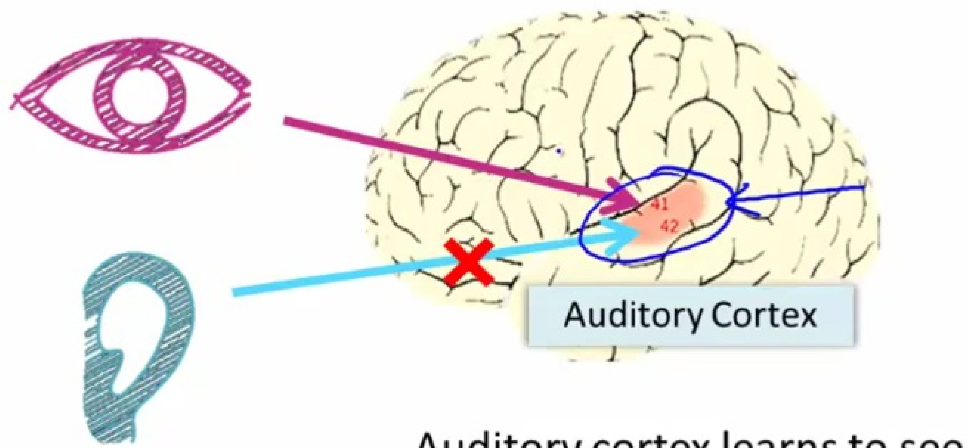
2.神经网络。历史：

Neural Networks

- Origins: Algorithms that try to mimic the brain.
- Was very widely used in 80s and early 90s; popularity diminished in late 90s.
- Recent resurgence: State-of-the-art technique for many applications

一个学习算法来模仿大脑假设：

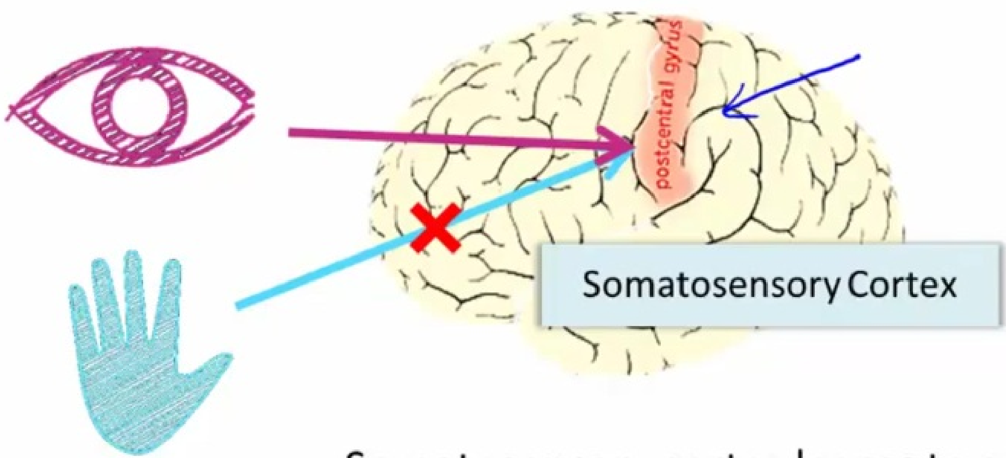
The “one learning algorithm” hypothesis



Auditory cortex learns to see

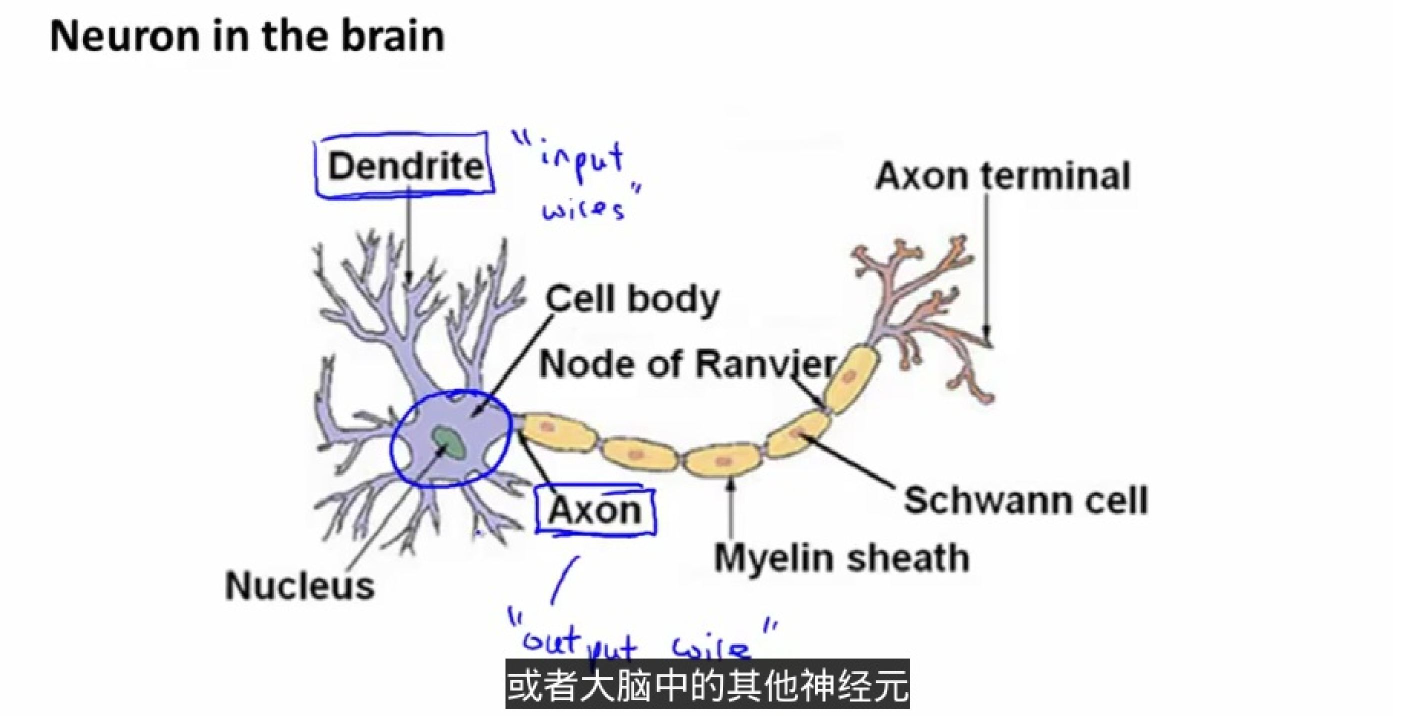
大脑里面处理听觉的部分转接到眼睛，就可以学会视觉。处理触觉的部分转接到眼睛也能完成视觉辨认任务。这称为神经重接实验。

The “one learning algorithm” hypothesis

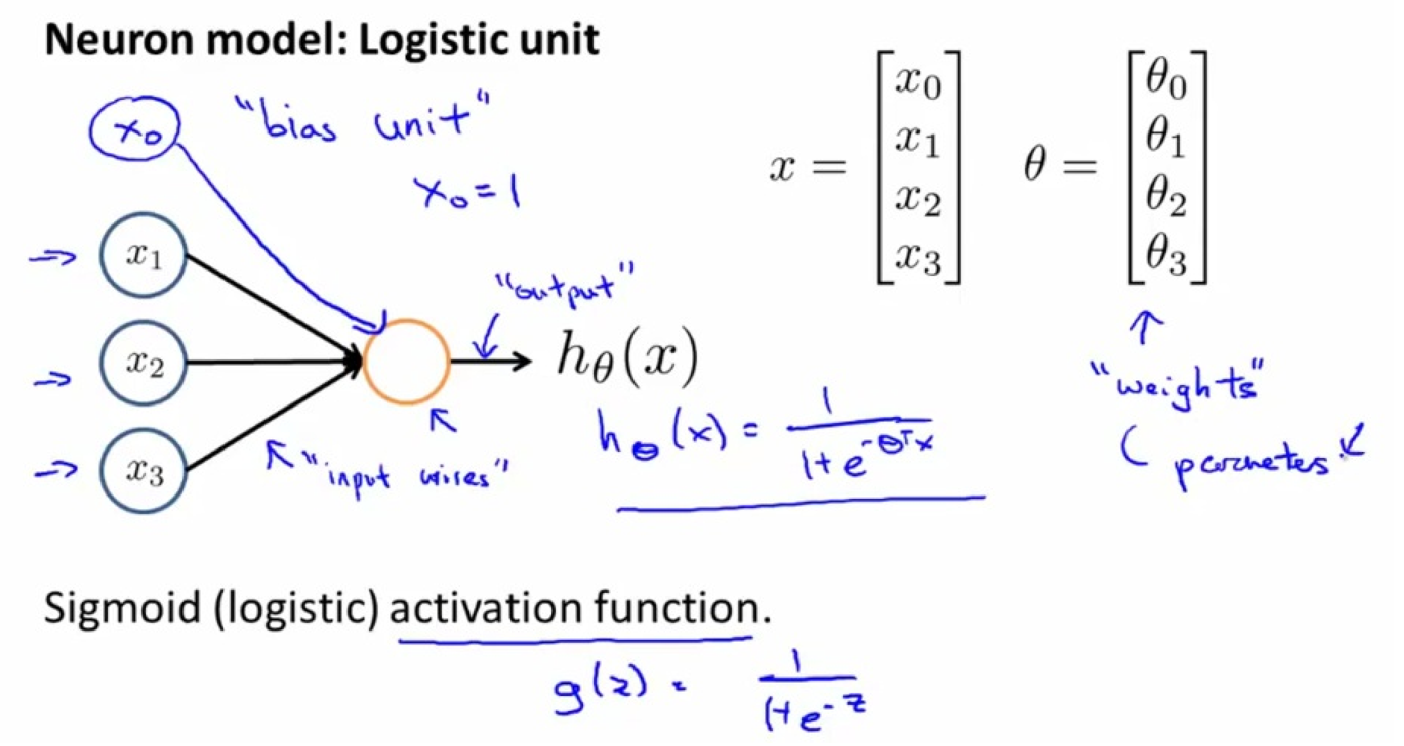


Somatosensory cortex learns to see

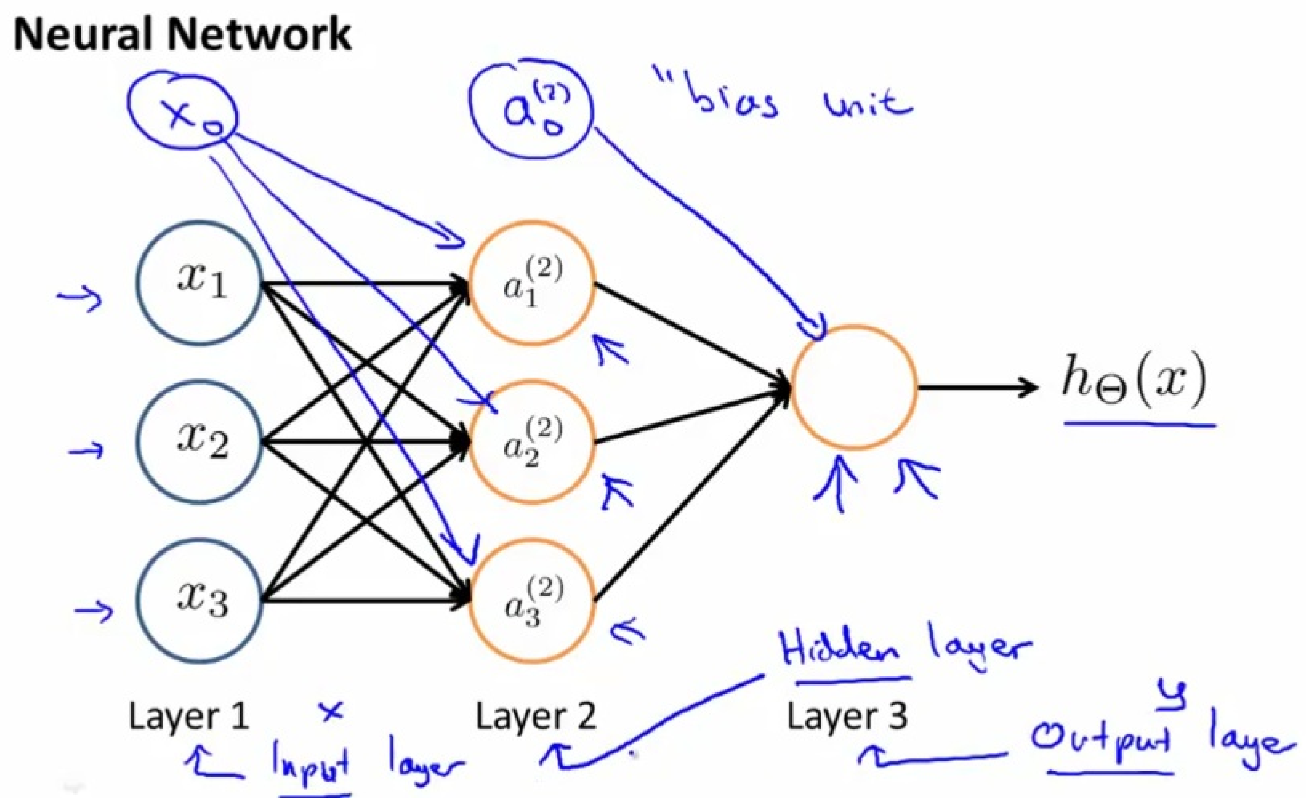
神经元结构：细胞，传入神经元，传出神经元。



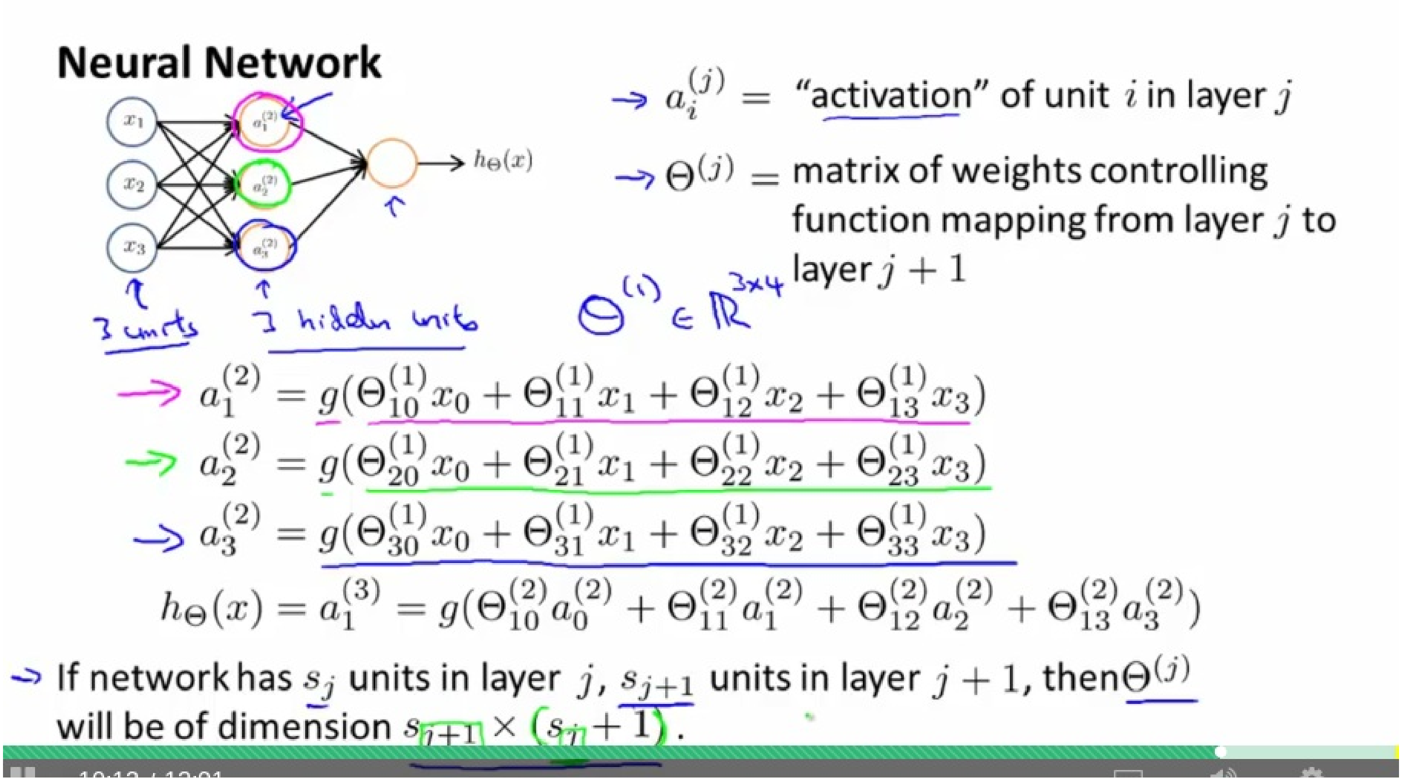
人工神经网络用逻辑单元来模拟神经元活动。（x0-偏置单元）下面是单一神经元：



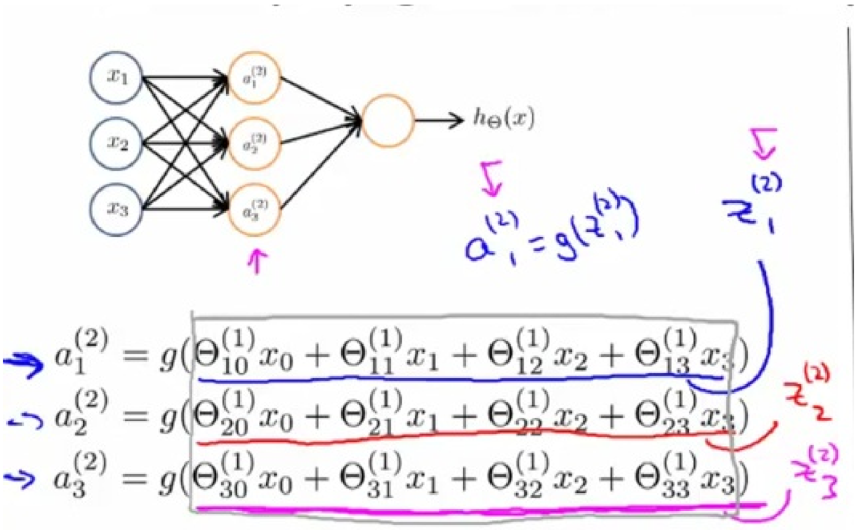
输入层，隐藏层，输出层：



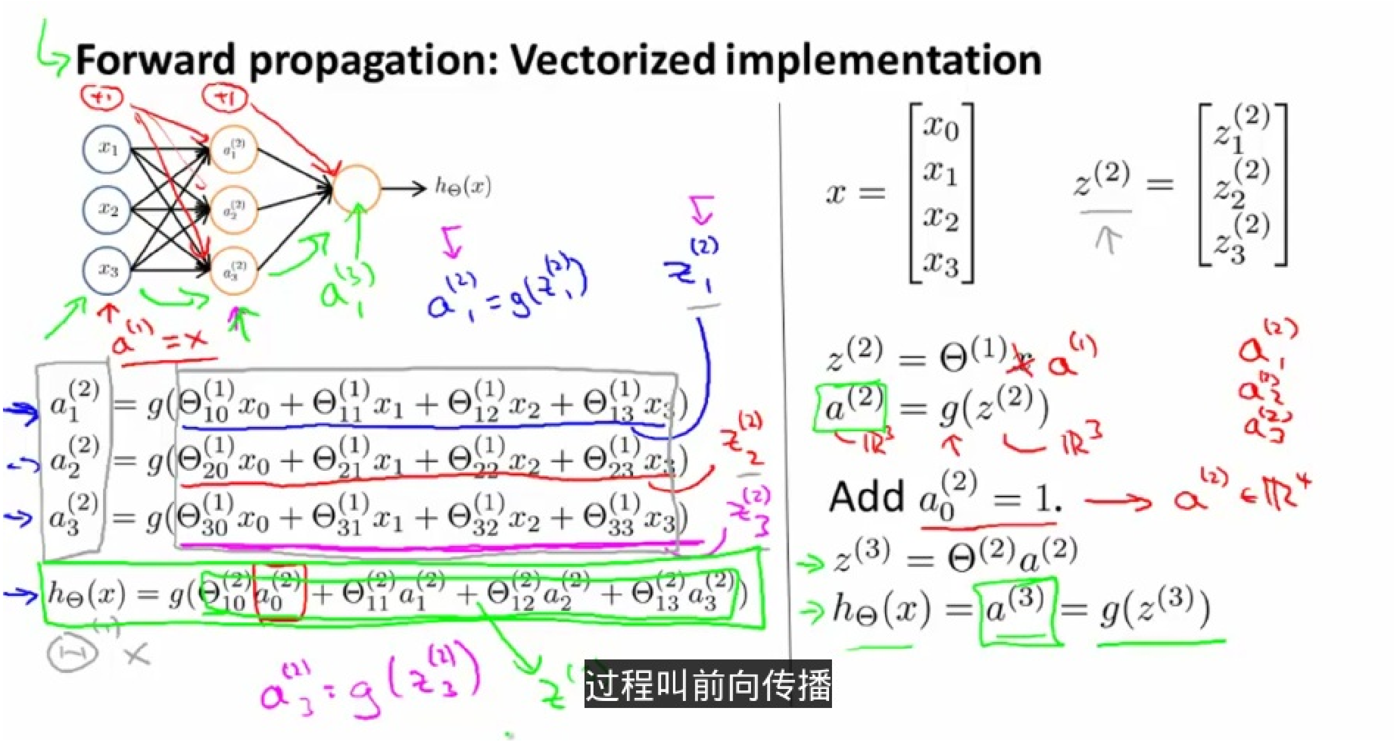
记号表示及映射矩阵的维度：



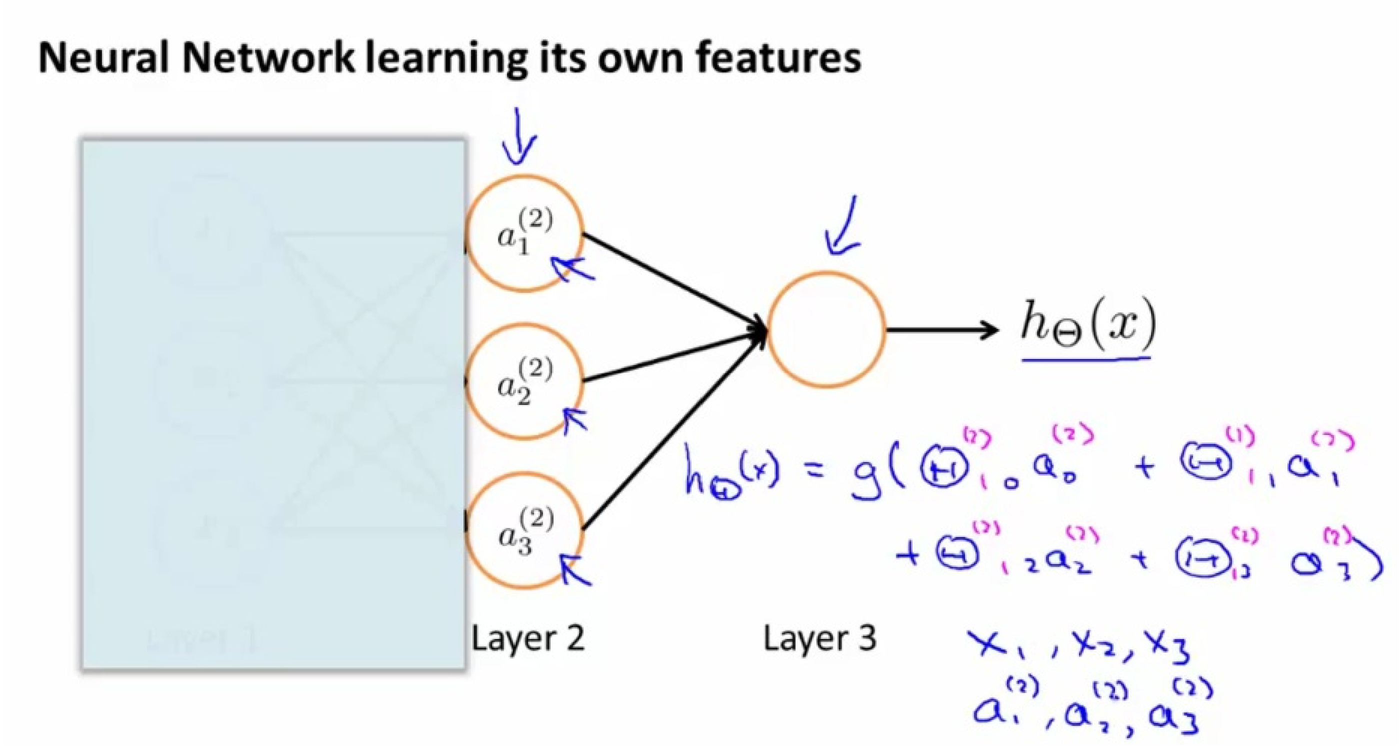
参数简化表示：z



向量化：->前向传播（Forward Propogation）

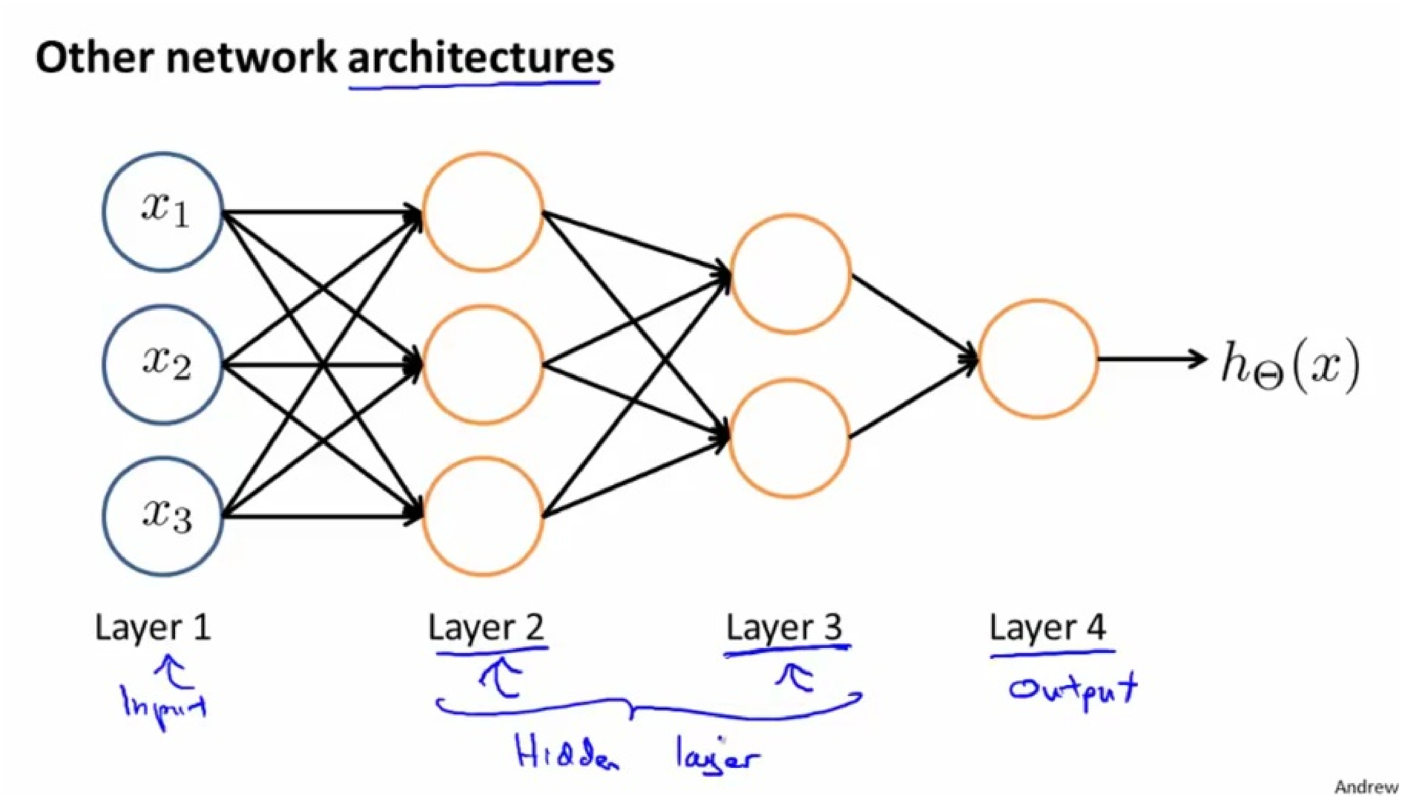


从某种角度看，神经网络就像是逻辑回归模型，只不过模型特性不是x,而是它自己的特性，即隐藏层：

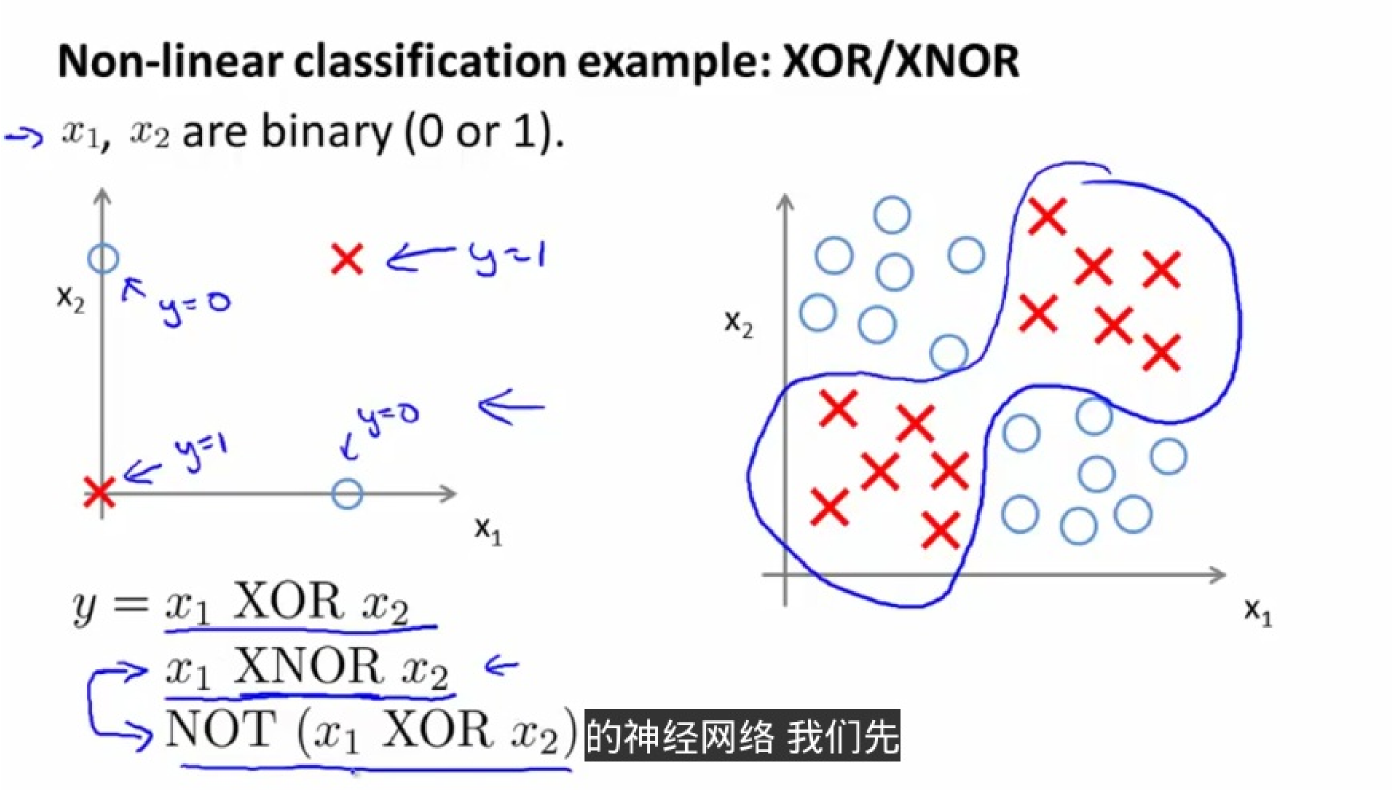


用不同的参数值(theta)来学习到隐藏层，然后把隐藏层作为特征来训练逻辑回归模型。

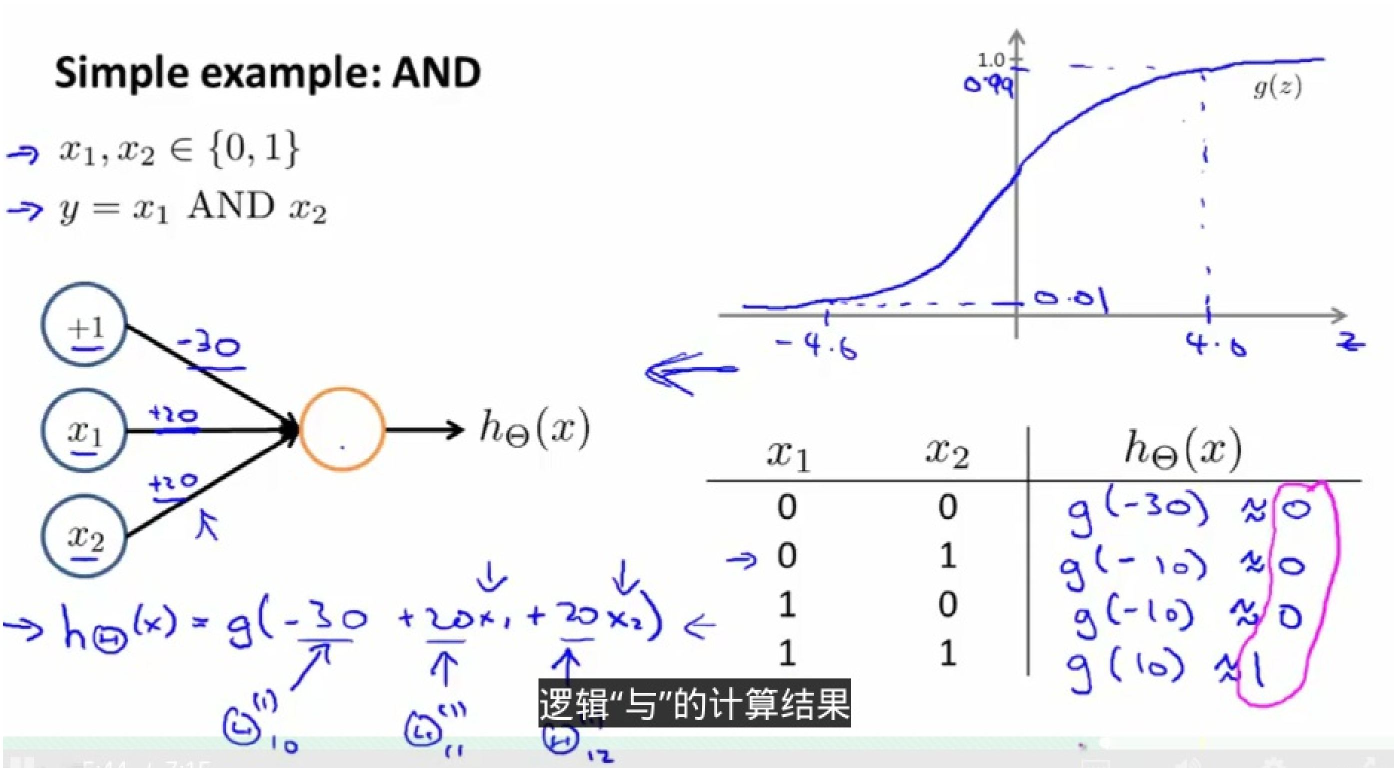
更复杂的网络架构：



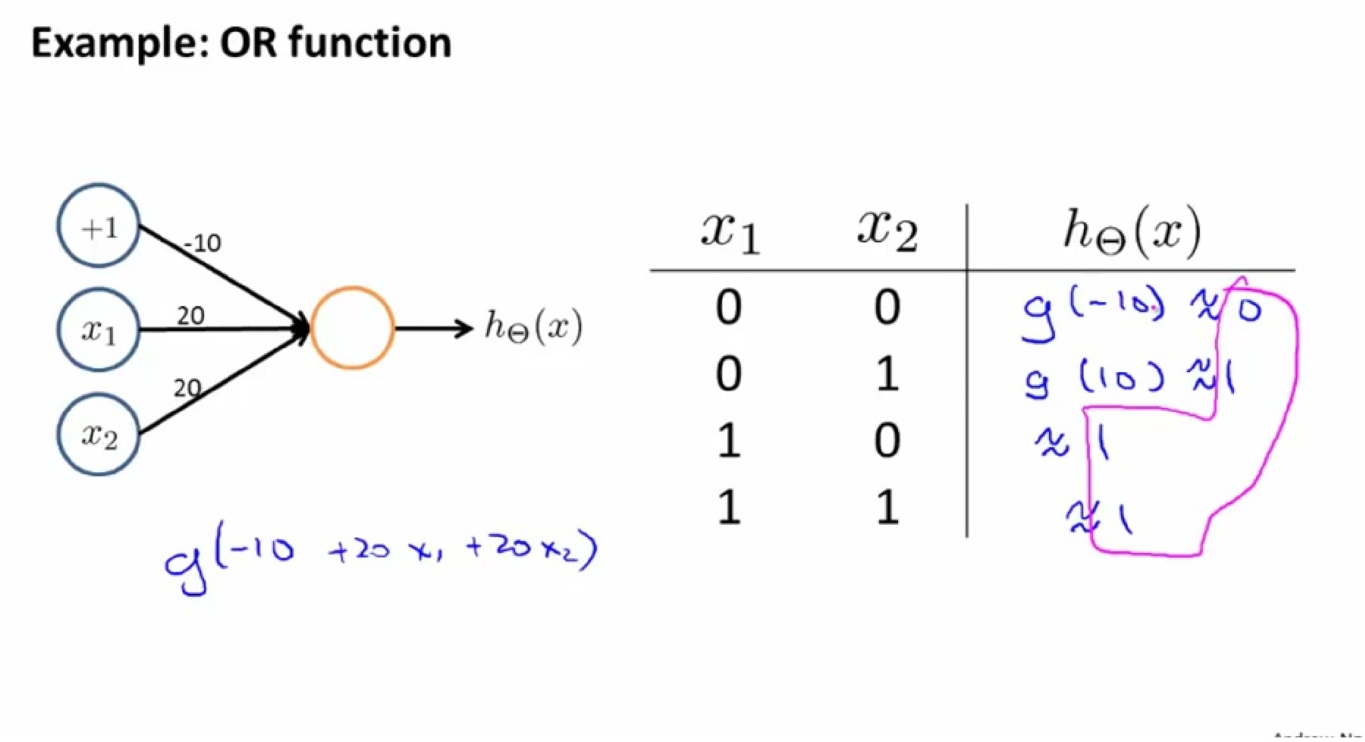
非线性分类例子：异或。



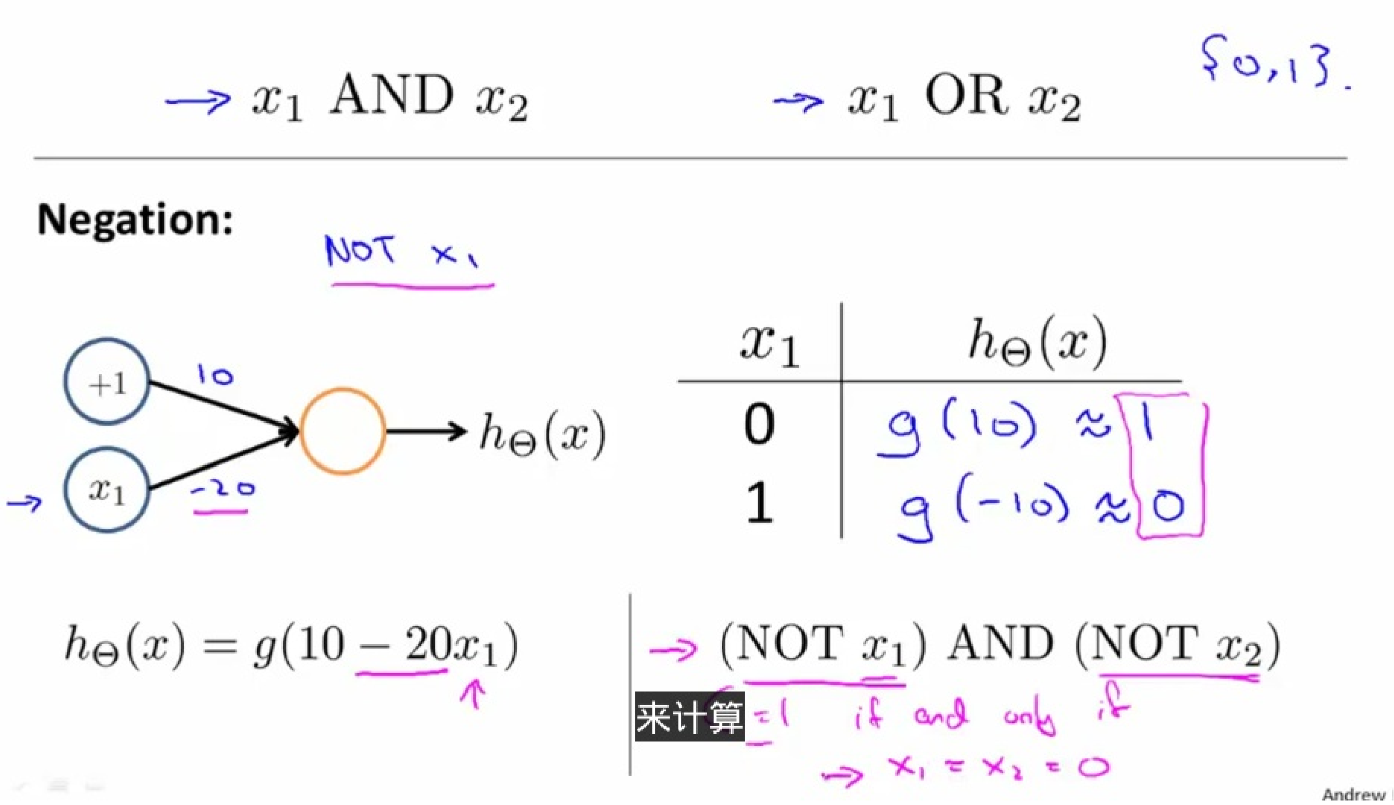
可以实现与运算的神经网络：



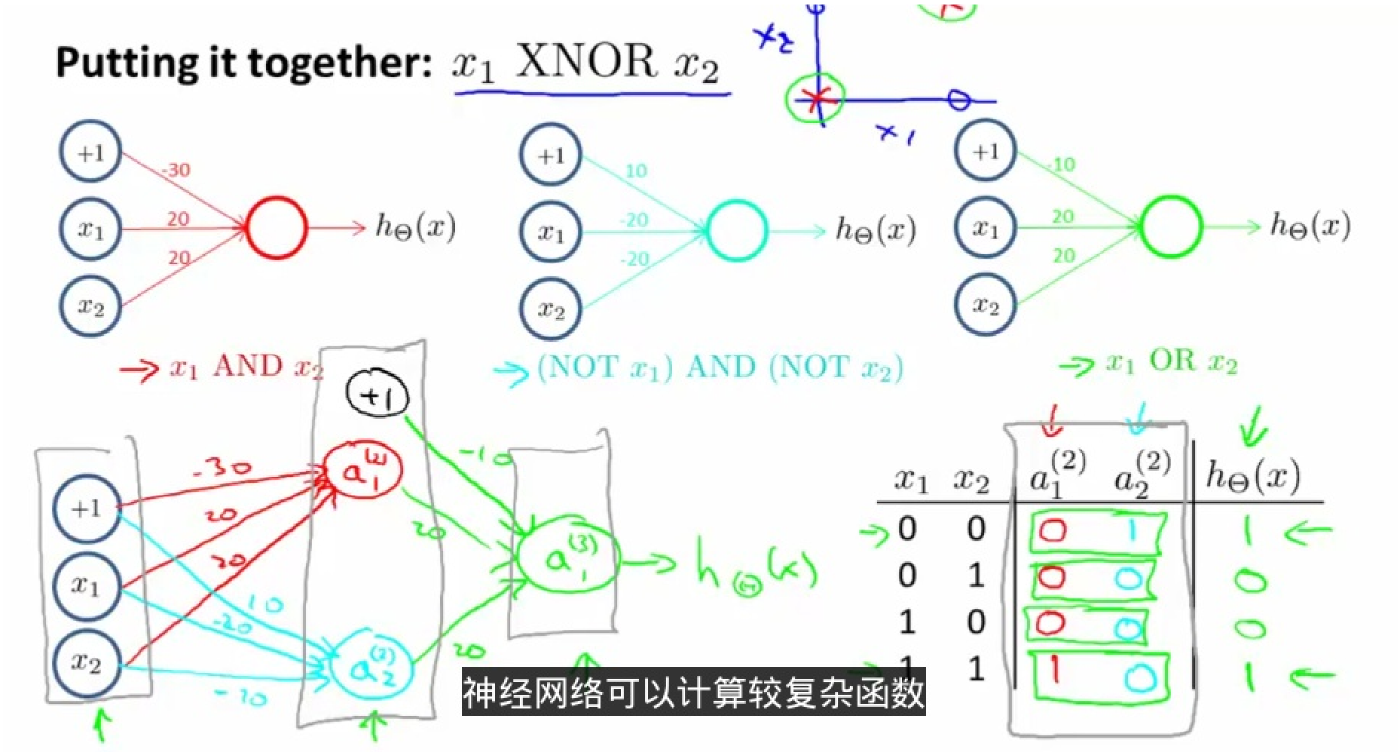
可以实现逻辑或的神经网络：



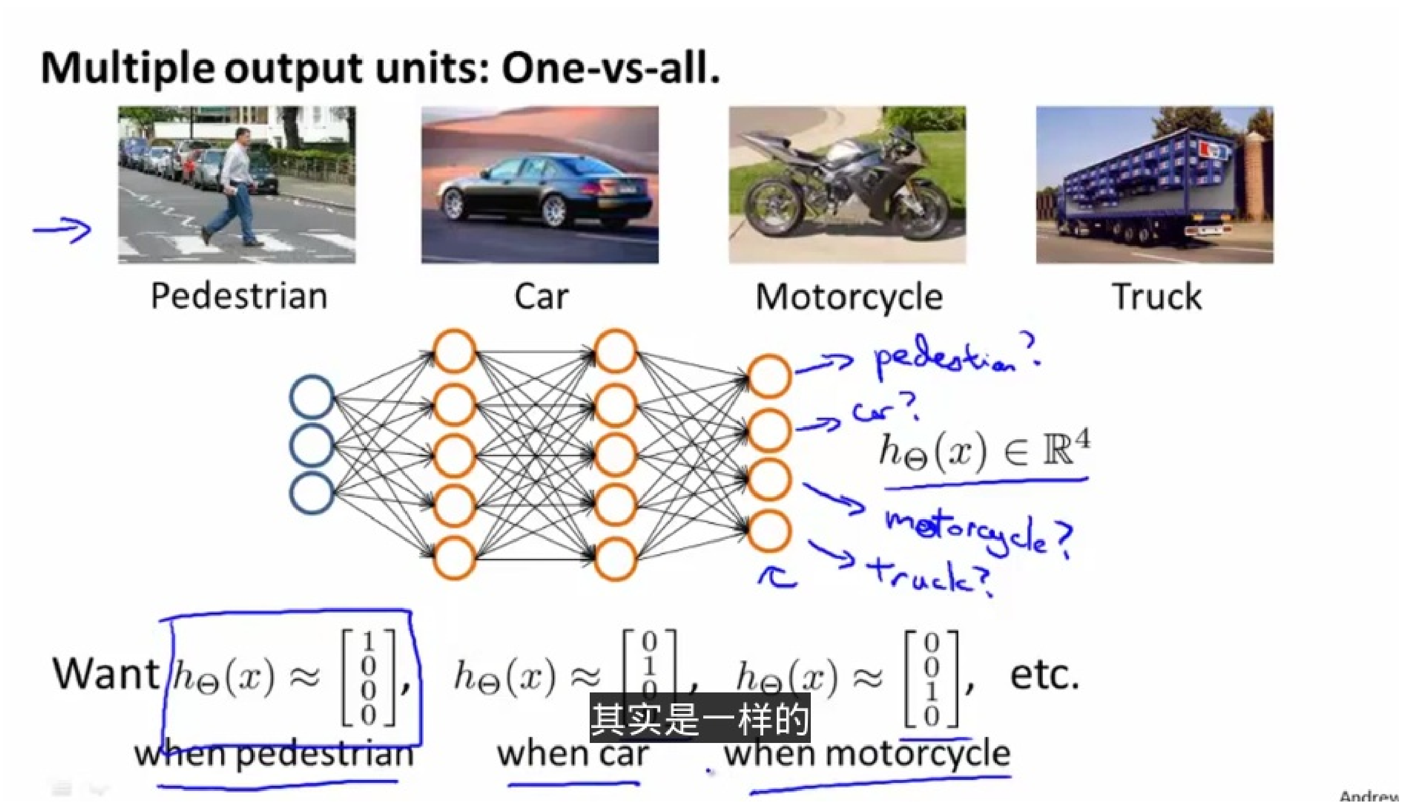
实现逻辑或的神经网络：



搭建人工神经网络实现异或：

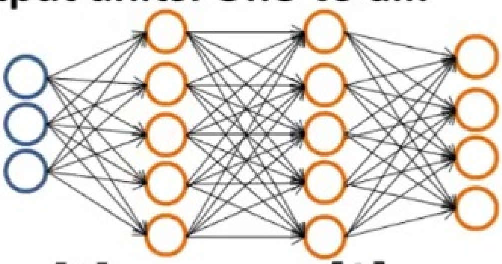


多输出单元：One-vs-All: 输出多元向量



训练方法：

Multiple output units: One-vs-all.



$h_\Theta(x) \in \mathbb{R}^4$

Want $h_\Theta(x) \approx \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $h_\Theta(x) \approx \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, etc.

when pedestrian   when car   when motorcycle

Training set: $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})$

~~Previously $y \in \{1, 2, 3, 4\}$~~

$y^{(i)}$ one of $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

pedestrian   car   motorcycle   truck

$(x^{(i)}, y^{(i)})$

$\underline{h_\Theta(x^{(i)})} \approx \underline{y^{(i)}}$

$\mathbb{R}^4$