

## 1.2 Deep Learning

Three Steps for deep learning



### Step1 Neural Network

有不同的方式连接这些 Neuron (手动连接)，最基本的是全连接前馈网络 (Full Connect Feedforward Network)

#### 1. 全连接 (Full Connect) :

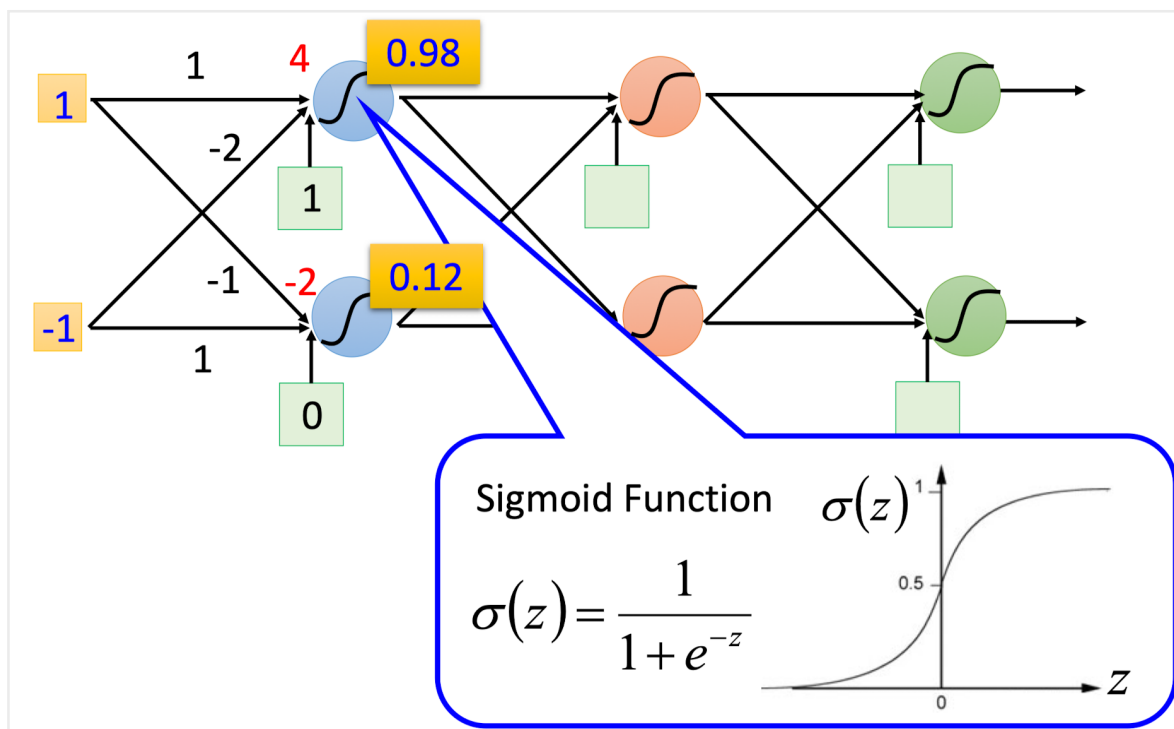
- 这意味着网络的每一层中的神经元与前一层和后一层中的所有神经元都连接在一起。
- 例如，第  $i$  层的第  $j$  个神经元接收来自第  $i - 1$  层的所有神经元的输入，并将其输出发送到第  $i + 1$  层的所有神经元。

#### 2. 前馈 (Feedforward) :

- 这意味着信息在网络中只有一个方向，即从输入层到输出层，不会有任何回路。这与某些其他类型的网络，如循环神经网络，形成了对比，后者允许信息从后面的层反馈到前面的层。

全连接前馈网络的结构通常由三种类型的层组成：

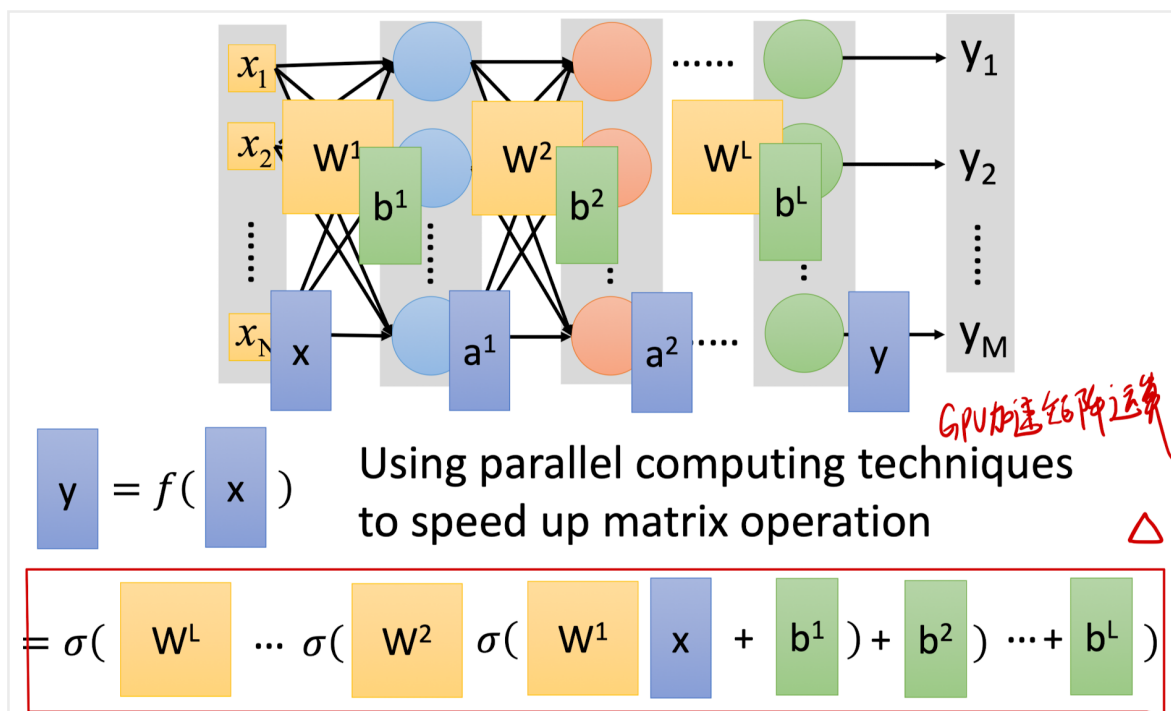
1. **输入层**：它直接接收数据输入。
2. **隐藏层**：位于输入层和输出层之间，可以有一个或多个隐藏层。
3. **输出层**：网络的最后一层，用于输出预测或分类结果。



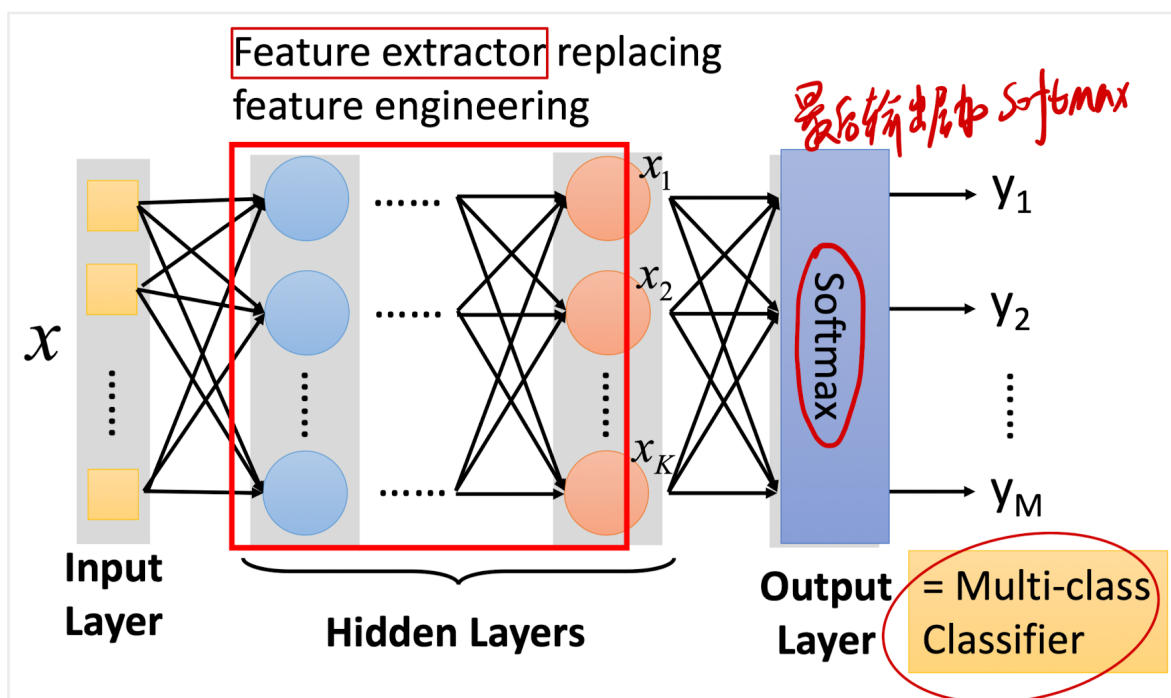
其实就是矩阵操作

$$\sigma\left( \underbrace{\begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\begin{bmatrix} 4 \\ -2 \end{bmatrix}} \right) = \begin{bmatrix} 0.98 \\ 0.12 \end{bmatrix}$$

整个neuron network的运算其实就是一连串的matrix operation(vectore \* matrix + vextor)，矩阵运算的好处是可以用GPU加速运算

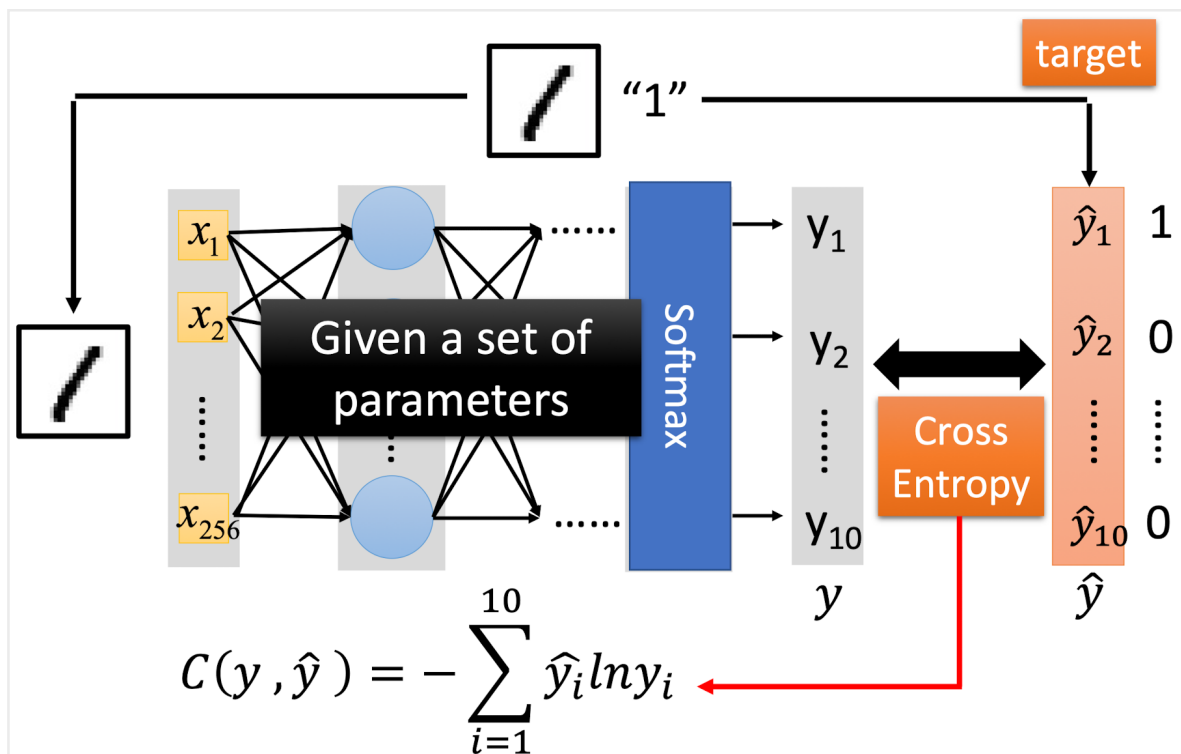


Hidden Layers 起的作用其实是 Feature extractor (特征提取器, feature engineering 是自己分析出合适的特征), Output Layer 其实是 Multi-class Classifier, 拿前一个 layer 的 output 当作 feature, 并通过一个 softmax function (Softmax 主要用于多分类问题的输出层, 但对于二分类问题, 使用 Sigmoid 激活)



## Step2 Goodness of function

计算  $y$  与  $y^{\text{head}}$  的 cross-entropy, 调整参数让 Total Loss 越小越好 (梯度下降)

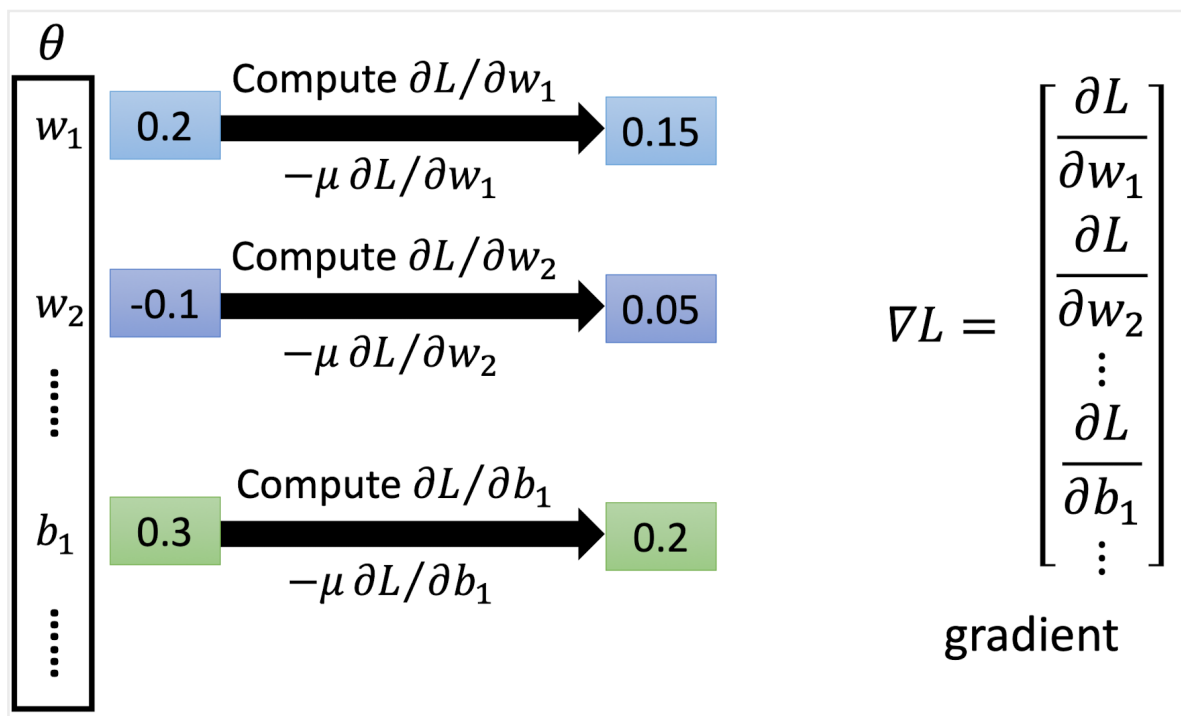


Total Loss:

$$L = \sum_{n=1}^N C^n$$

Step3 pick the best function

Gradient Descent



### Backpropagation

- Backpropagation: an efficient way to compute  $\partial L / \partial w$  in neural network