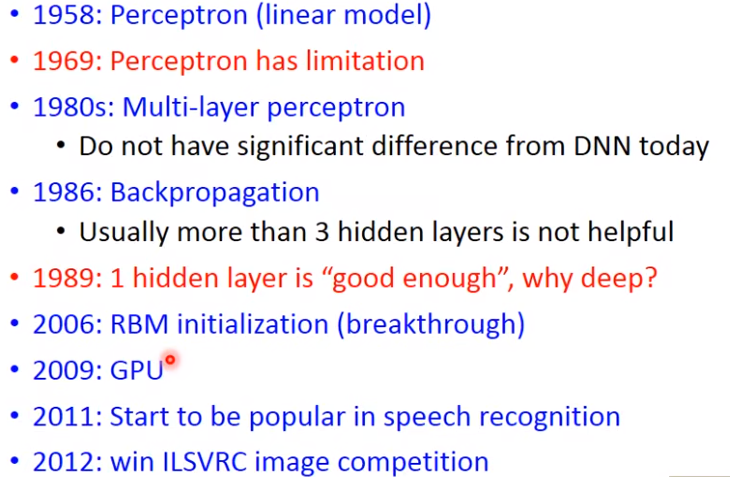
# 机器学习笔记

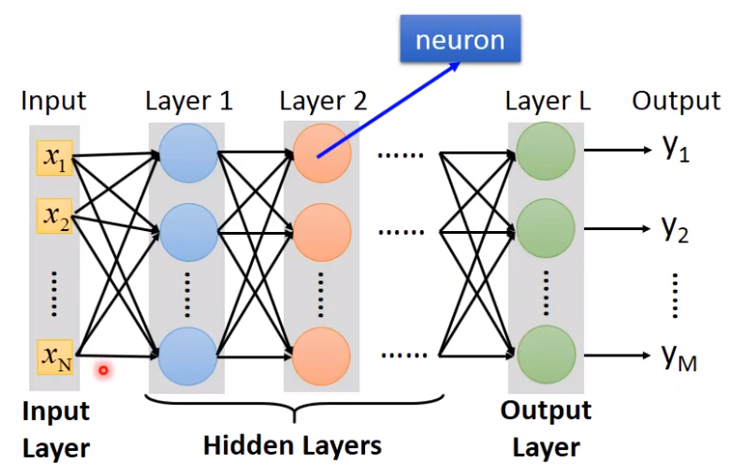
李宏毅教授机器学习(视频地址：https://www.bilibili.com/video/BV1JE411g7XF?p=3&vd\_source=0a27b8c0284e7cd3b72bc0c89478f2d7)

# 第五课：Brief Introduction of Deep Learning、

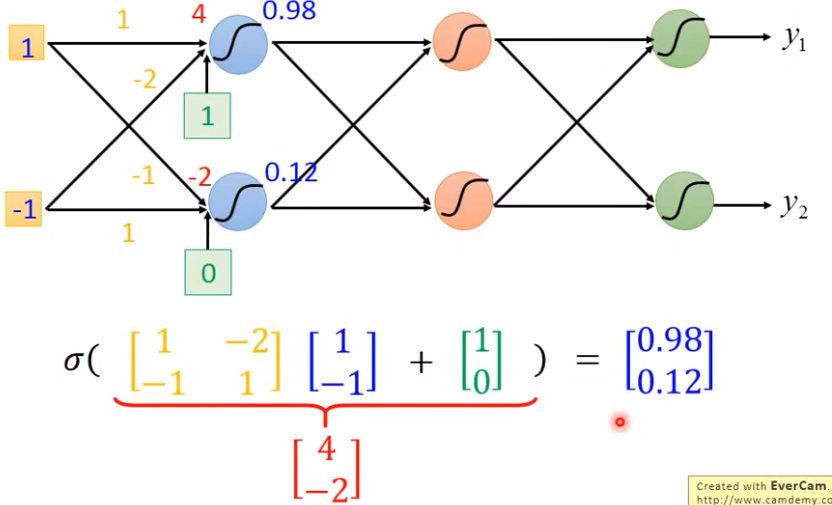
Deep Learning的简要发展史如图所示，



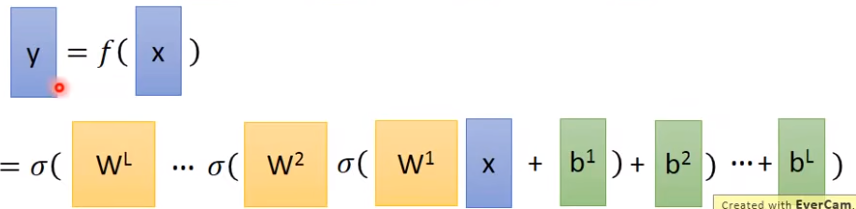
Deep Learning的步骤和Machine Learning的步骤几乎是一样的，都是先定义一些function，然后评估这些function的好坏程度，最后选出一个最好的function。而这里得出来的function其实是一个neural network(神经网络)，在这个神经网络里面我们有很多的logistic regression，我们需要将它们连接起来，而这个连接的方法有很多种，最简单的一种就是Fully Connected Feedforward Network，这种做法是把neuron排成一排一排，然后将它们完全连接起来，相邻两排neuron可以粗略地看作是完全二分图，对于不同层来说，它们有不同的称呼，如下图，最左层称为Input Layer，最右层称为Output Layer，而剩下中间的都称为Hidden Layer。



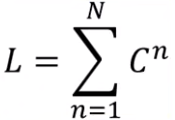
不同层之间的输入输出关系我们可以看作是矩阵运算，如下图所示，层之间的关系用一个矩阵表示，输入用一个列向量表示，通过矩阵运算之后得到的输出就作为下一层的输入。

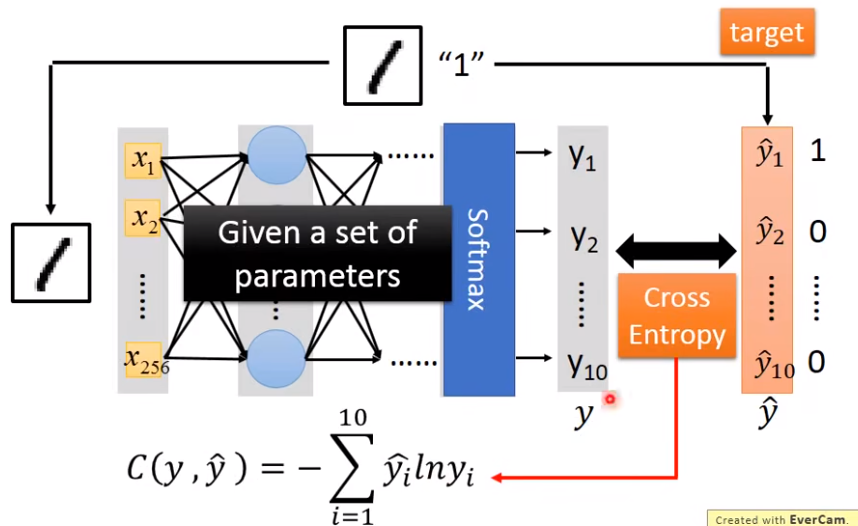


我们把每一层的w和b集合起来当作一个W和B，然后输入x集合起来当作X，那么神经网络做的事就如图所示，像“套娃”一样的过程，



我们最后Output Layer要做的事就是将前一个layer的Output当作feature然后经过softmax function的转换进行输出。

我们定义的Deep Learning中某一个function的好坏程度是用loss function来评估，如下图所示，我们想要的目标输出，而其余的都为0，然后我们用最后Output Layer的输入y与之间做Cross Entropy，然后loss function的公式就写成，



令loss function最小化的方法就是之前用到的Gradient Descent。