上一篇笔记《机器学习：Deep Learning》中简单介绍了前馈神经网络的结构、矩阵操作和简单应用，本篇文章则继续介绍Deep Learning相关训练的Loss function和参数计算问题。

**1 Loss function**

当Neural个数和连接方式确以后，神经网络结构就确定了，即模型已经确定，接下来就要确定一个策略，来保证模型是最优的，一般我们就定义一个Loss function。如图1所示：



图1 Loss function定义

如图1所示，Loss function定义为Cross Entropy，训练序列为，：

单个训练数据的Loss function：



总的训练数据的Loss function：



最后就是优化参数，使得最小，在Deep Learning优化参数的方法主要还是梯度下降法。







**2 Back propagation**

在Deep Learning中随着Hidden Layer的增加，参数也越来越多，利用梯度下降法计算参数的时候就要耗费巨大的计算资源，如何能高效率的计算参数是很重要的，Back propagation是一种高效算法。

假设一个Deep Learning的结构如图2所示：



图2\_1 Neuron



图2\_2 Deep Learning示例

如果按正常的算法会比较浪费计算资源，因为在顺序迭代的过程中一开始我们并不知道是多少，需要我们先对前向网络进行一轮计算，得出再又返回来对前向网络再计算一轮，这个过程中很多求导或者计算是重复的。

现在假设求参数的最优解，则：



1）forward pass就是对所有的计算；

2）backward pass就是对所有的计算。

forward pass的计算是比较简单的，因为我们知道，的结果就是或者等等。

backward pass的计算：



因为实际上是对函数求导数：；而这个步骤在forward pass中就可以完成，所以关键在于如何求：



根据图2\_2可知：



则：



实际上也是已知的，因为在forward pass中就被计算出来。而计算和和计算一样，这个过程就是backward pass的计算。

总体来说，**Back propagation**的计算分以下两个步骤，如图3所示：



图3 Back propagation计算示意图



2017.08.19.