# 实验八 前馈神经网络（4）

班级： 学号： 姓名：

# 实验任务：优化问题

**注意事项：不要粘贴源代码。实验报告中只需描述算法和操作流程，分析实验结果。**

1. 基本概念

**1 参数初始化**

实现一个神经网络前，需要先初始化模型参数。如果对每一层的权重和偏置都用0初始化，那么通过第一遍前向计算，所有隐藏层神经元的激活值都相同；在反向传播时，所有权重的更新也都相同，这样会导致隐藏层神经元没有差异性，出现**对称权重现象**。

**2 梯度消失问题**

在神经网络的构建过程中，随着网络层数的增加，理论上网络的拟合能力也应该是越来越好的。但是随着网络变深，参数学习更加困难，容易出现梯度消失问题。

由于Sigmoid型函数的饱和性，饱和区的导数更接近于0，误差经过每一层传递都会不断衰减。当网络层数很深时，梯度就会不停衰减，甚至消失，使得整个网络很难训练，这就是所谓的**梯度消失问题**。

在深度神经网络中，减轻梯度消失问题的方法有很多种，一种简单有效的方式就是使用导数比较大的激活函数，如：ReLU。

**3 死亡ReLU问题**

ReLU激活函数可以一定程度上改善梯度消失问题，但是在某些情况下容易出现**死亡ReLU问题**，使得网络难以训练。

这是由于当x<0时，ReLU函数的输出恒为0。在训练过程中，如果参数在一次不恰当的更新后，某个ReLU神经元在所有训练数据上都不能被激活（即输出为0），那么这个神经元自身参数的梯度永远都会是0，在以后的训练过程中永远都不能被激活。

一种简单有效的优化方式就是将激活函数更换为Leaky ReLU、ELU等ReLU的变种。

1. 编码实现

根据实验教材代码，使用Pytorch完成相应程序编写。

#### 实验结果：

XXX

# 实验总结：

XXX

REF：《神经网络与深度学习：案例与实践》