

第五周

Day1-Day2

注意：回顾题一般为课本上的题目，需要用自己的话进行总结打卡。补充题则为课本上没有另行补充的题目，会提供一些参考答案。

1. 回顾：画出在二元输入的情况下，表示异或逻辑的网络图

参考课本9.1节进行总结

2. 回顾：写出Sigmoid、tanh、relu的激活函数和对应的导数

参考课本9.2节进行总结

3. 回顾：为什么引入非线性激励函数？

因为如果不用非线性激励函数，每一层都是上一层的线性函数，无论神经网络多少层，输出都是输入的线性组合，与只有一个隐藏层效果一样。相当于多层感知机了。所以引入非线性激励函数，深层网络就变得有意义了，可以逼近任意函数。

4. 回顾：说明神经网络参数是否可以初始化为同一个值，为什么？

参考课本9.4节进行总结

5. 回顾：说明dropout抑制过拟合的原理

参考课本9.4节进行总结

6. 回顾：说明为什么需要批量归一化

参照课本9.4节进行总结

6. 补充：CNN出现梯度消失怎么办？

过多的层数会导致梯度消失，解决手段有减少层数；增大学习率；用Relu代替sigmoid

7. 回顾：dropout内部是怎么实现只让部分信号通过并不更新其余部分

以前我们网络的计算公式是：

$$z_i^{(l+1)} = w_i^{(l+1)} y^l + b_i^{(l+1)},$$

$$y_i^{(l+1)} = f(z_i^{(l+1)}).$$

采用dropout后计算公式就变成了：

$$r_j^{(l)} \sim \text{Bernoulli}(p),$$

$$\tilde{y}^{(l)} = r^{(l)} * y^{(l)},$$

$$z_i^{(l+1)} = w_i^{(l+1)} \tilde{y}^l + b_i^{(l+1)},$$

$$y_i^{(l+1)} = f(z_i^{(l+1)}).$$

上面公式中Bernoulli函数，是为了以概率p，随机生成一个0、1的向量。

8. 补充：增加神经网络的层数有什么用

深度神经网络的成功，已经证明，增加神经网络层数，可以增加模型泛化能力，即，训练数据集和测试数据集都表现得更好。但是更多的层数，不一定能保证有更好的表现。

9. 补充：为什么CNN要用权值共享

减少参数数量，使运算变得简洁、高效，能够在超大规模数据集上运算。

10. 补充：CNN的局部连接是什么

谓局部连接，就是卷积层的节点仅仅和其前一层的部分节点相连接，只用来学习局部特征。

在图像中，距离较近的像素间相关性强，距离较远则相关性就比较弱，图像的特征是局部强相关。因此，局部感知采用部分神经元接受图像信息，再通过综合全部的图像信息达到增强图像信息的目的。

局部连接的方式大幅减少了参数数量，加快了学习速率，同时也在一定程度上减少了过拟合的可能。