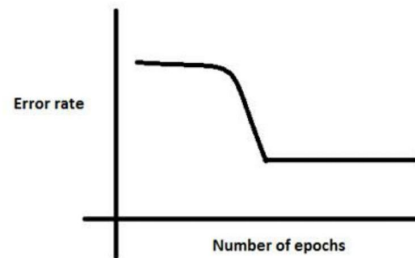


1. 在训练神经网络时，损失函数（Loss）在最初的几个 epochs 时没有下降，可能的原因是？



学习率太低；正则参数太高；陷入局部最小值

2. 批规范化（Batch Normalization）的好处有？

- a. 让每一层的输入的范围都大致固定
- b. 它将权重的归一化平均值和标准差
- c. 它是一种非常有效的反向传播（BP）方法
- d. 这些均不是

a

3. 在一个神经网络中，下面哪种方法可以用来处理过拟合？

- a. Dropout
- b. Batch Normalization
- c. 正则化
- d. 以上都可以

d

4. 考虑某个具体问题，你可能只有少量数据来解决这个问题。不过幸运的是你有一个类似问题已经预先训练好的神经网络。可以用下面哪种方法来利用这个预先训练好的网络？

- a. 把除了最后一层外所有的层都冻结，重新训练最后一层
- b. 对新数据重新训练整个模型
- c. 只对最后几层进行调参（fine tune）
- d. 对每一层模型进行评估，选择其中的少数来使用

c

5. 下图是一个利用 sigmoid 函数作为激活函数的含四个隐藏层的神经网络训练的梯度下降图。这个神经网络遇到了梯度消失的问题。下面哪个叙述是正确的（纵坐标表示梯度大小）？