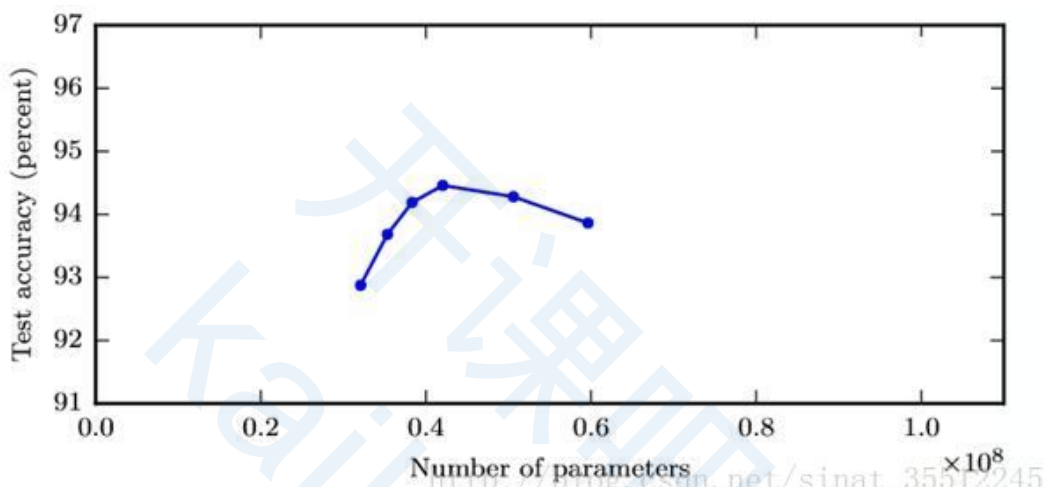


1. 卷积神经网络可以对一个输入进行多种变换（旋转、平移、缩放），这个表述正确吗？

2. 下列哪一项在神经网络中引入了非线性？（ ）

- A. 随机梯度下降
- B. 修正线性单元（ReLU）
- C. 卷积函数
- D. 以上都不正确

3. 下图显示了训练过的 3 层卷积神经网络准确度，与参数数量(特征核的数量)的关系。



从图中趋势可见，如果增加神经网络的宽度，精确度会增加到一个特定阈值后，便开始降低。造成这一现象的可能原因是什么？（ ）

- A. 即使增加卷积核的数量，只有少部分的核会被用作预测
- B. 当卷积核数量增加时，神经网络的预测能力（Power）会降低
- C. 当卷积核数量增加时，它们之间的相关性增加(correlate)，导致过拟合
- D. 以上都不正确

4. 假设我们已经在 ImageNet 数据集(物体识别)上训练好了一个卷积神经网络。然后给这张卷积神经网络输入一张全白的图片。对于这个输入的输出结果为任何种类的物体的可能性都是一样的，对吗？（ ）

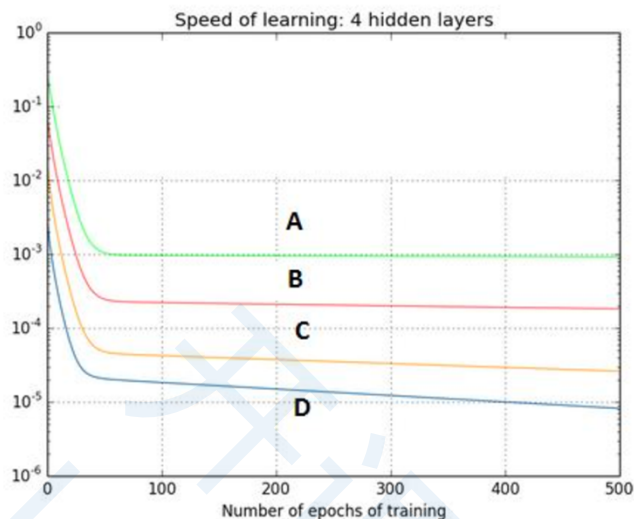
- A. 对的
- B. 不知道
- C. 看情况
- D. 不对

5. 当在卷积神经网络中加入池化层(pooling layer)时，变换的不变性会被保留，是吗？（ ）

- A. 不知道
- B. 看情况
- C. 是

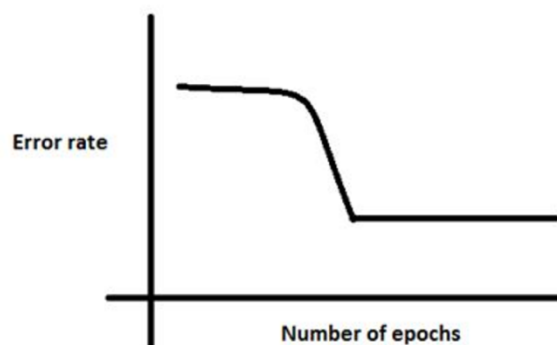
D. 否

6. 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？（ ）



http://blog.csdn.net/sinat_35512245

- A. 第一隐藏层对应 D，第二隐藏层对应 C，第三隐藏层对应 B，第四隐藏层对应 A
- B. 第一隐藏层对应 A，第二隐藏层对应 C，第三隐藏层对应 B，第四隐藏层对应 D
- C. 第一隐藏层对应 A，第二隐藏层对应 B，第三隐藏层对应 C，第四隐藏层对应 D
- D. 第一隐藏层对应 B，第二隐藏层对应 D，第三隐藏层对应 C，第四隐藏层对应 A
7. 下图显示，当开始训练时，误差一直很高，这是因为神经网络在往全局最小值前进之前一直被卡在局部最小值里。为了避免这种情况，我们可以采取下面哪种策略？（ ）



http://blog.csdn.net/sinat_35512245

- A. 改变学习速率，比如一开始的几个训练周期不断更改学习速率

- B. 一开始将学习速率减小 10 倍，然后用动量项(momentum)
C. 增加参数数目，这样神经网络就不会卡在局部最优处
D. 其他都不对
8. 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住，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，选择其中的少数来用
9. 增加卷积核的大小对于改进卷积神经网络的效果是必要的吗？
10. 输入图片大小为 200×200 ，依次经过一层卷积（kernel size 5×5 ，padding 1，stride 2），pooling（kernel size 3×3 ，padding 0，stride 1），又一层卷积（kernel size 3×3 ，padding 1，stride 1）之后，输出特征图大小为（ ）：
A. 95
B. 96
C. 97
D. 98
11. 深度学习是当前很热门的机器学习算法，在深度学习中，涉及到大量的矩阵相乘，现在需要计算三个稠密矩阵 A,B,C 的乘积 ABC,假设三个矩阵的尺寸分别为 $m \times n$ ， $n \times p$ ， $p \times q$ ，且 $m < n < p < q$ ，以下计算顺序效率最高的是（ ）
A. (AB)C
B. AC(B)
C. A(BC)
D. 所以效率都相同
12. 卷积网络有哪些池化操作，其作用又是什么？
13. Dropout 具体做了什么使得其能够降低模型的过拟合？
14. 请解释下 CNN 网络中权值共享与稀疏交互的特性及作用？