

作业二

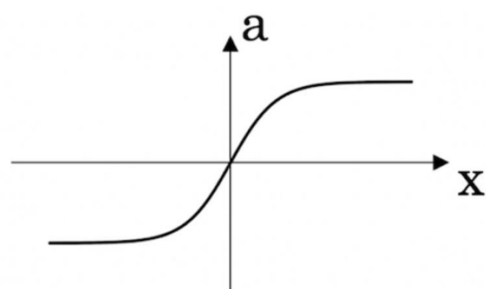
一. 判断题

1. 神经网络随着网络层数的增加表达能力变得更强, 因此网络中的激活函数可以去掉, 只需要加深网络就能够训练足够强的神经网络。()
2. Long short-term Memory (LSTM)网络具有记忆和遗忘功能, 适用于序列建模。()
3. 注意力机制(attention)相比 LSTM 更高效, 因为在每个 time step 计算时都可以读到全局信息, 而不需要像 LSTM 那样串行计算。()
4. 训练卷积神经网络 (CNN) 时, 如果对训练样本通过平移、旋转和缩放等操作额外生成一些补充样本, 会从整体上降低训练样本的质量, 影响网络提取特征, 从而导致预测准确率下降。()
5. 在网络中加入 Dropout 和 Batch Normalization 都是深度学习中常见的防止过拟合的手段, 在所有类型的深度神经网络中都适用。()

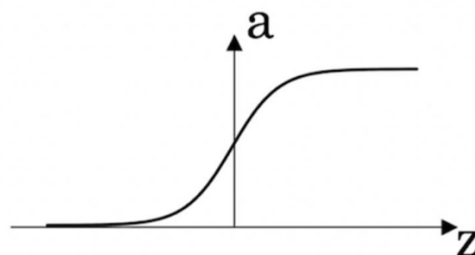
二. 选择题

1. 下图中哪一个表述 ReLU 激活函数? ()

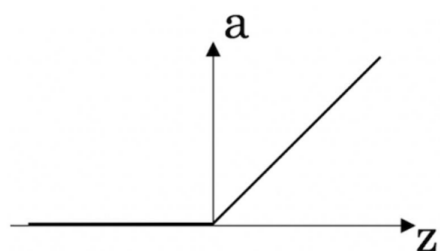
A.



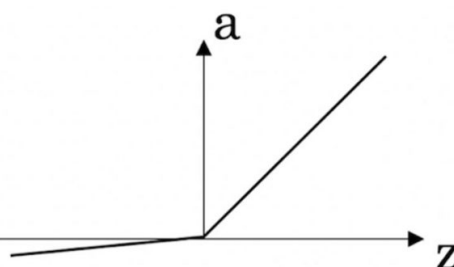
B.



C.



D.



2. 输入为 64×64 的 RGB 图片, 使用 32 个 3×3 的卷积核提取特征, 步长为 1, 不使用 padding 填充, 则输出的大小为: ()
(a) $64 \times 64 \times 64$ (b) $64 \times 32 \times 32$ (c) $32 \times 62 \times 62$ (d) $32 \times 64 \times 64$
3. 上一题中, 该卷积层的参数有多少个? ()
(a) 64 (b) 32 (c) $32 \times 3 \times 3$ (d) $32 \times 3 \times 3 \times 3$

三 . 简答题

1. 深度学习模型为何在训练中容易出现过拟合？试分析原因并给出如何在深度学习模型训练中缓解过拟合？
2. 深度学习模型训练过程中为何会出现梯度消失和梯度爆炸问题？有哪些方法可以解决梯度消失或梯度爆炸？