

1. 对于一个图像识别问题(在一张照片里找出一只猫), 下面哪种神经网络可以更好地解决这个问题? ()

- A. 循环神经网络
- B. 感知机
- C. 多层感知机
- D. 卷积神经网络

解析:

卷积神经网络将更好地适用于图像相关问题, 因为考虑到图像附近位置变化的固有性质。

答案: (D)

2. (判断题) 在深层网络或循环神经网络中, 误差梯度可在更新中累积, 变成非常大的梯度, 然后导致网络权重的大幅更新, 并因此使网络变得不稳定。在极端情况下, 权重的值变得非常大, 以至于溢出, 导致 NaN 值。()

解析: 正确

3. 以下是一些稍微明显一点的信号, 哪儿些是有助于确认是否出现梯度爆炸问题的。()

- A. 训练过程中模型梯度快速变大。
- B. 训练过程中模型权重变成 NaN 值。
- C. 训练过程中, 每个节点和层的误差梯度值持续超过 1.0。
- D. 以上都是

解析: D

4. 关于修复梯度爆炸问题, 以下说法正确的是 ()

- A. 在循环神经网络中, 梯度爆炸的发生可能是因为某种网络的训练本身就存在不稳定性, 如随时间的反向传播本质上将循环网络转换成深度多层感知机神经网络。
- B. 使用长短期记忆 (LSTM) 单元和相关的门类型神经元结构可以减少梯度爆炸问题。
- C. 处理梯度爆炸有一个简单有效的解决方案: 如果梯度超过阈值, 就截断它们。
- D. 在深度多层感知机神经网络中, 梯度爆炸的发生可能是因为激活函数, 如之前很流行的 Sigmoid 和 Tanh 函数

解析: ABCD

5. LSTM 为什么比 RNN 好?

解析:

推导 forget gate, input gate, cell state, hidden information 等的变化; 因为 LSTM 有进有出且当前的 cell informaton 是通过 input gate 控制之后叠加的, RNN 是叠乘, 因此 LSTM 可以防止梯度消失或者爆炸。

6. 你有哪些 Deep Learning (RNN, CNN) 调参的经验?

解析:

答案解析, 来自知乎

7. 什么是 RNN?

解析:

循环神经网络(RNN, Recurrent Neural Networks)介绍

8. RNN 是怎么从单层网络一步一步构造的?

解析:

完全图解 RNN、RNN 变体、Seq2Seq、Attention 机制

9. 如何解决 RNN 梯度爆炸和弥散的问题的?

解析:

深度学习与自然语言处理(7)_斯坦福 cs224d 语言模型, RNN, LSTM 与 GRU

10. 为什么 LSTM 模型中既存在 Sigmoid 又存在 Tanh 两种激活函数?

解析:

本题解析来源: 知乎解答

二者目的不一样: sigmoid 用在了各种 gate 上, 产生 0~1 之间的值, 这个一般只有 sigmoid 最直接了。
tanh 用在了状态和输出上, 是对数据的处理, 这个用其他激活函数或许也可以。另可参见 A Critical Review of Recurrent Neural Networks for Sequence Learning 的 section4.1, 说了那两个 tanh 都可以替换成别的。

11. (判断) 自动编码器是一种数据的压缩算法, 其中数据的压缩和解压缩函数是数据相关的、有损的、从样本中自动学习的。在大部分提到自动编码器的场合, 压缩和解压缩的函数是通过神经网络实现的。()

解析: 正确

深度学习之自编码器 AutoEncoder