



六方云算法工程师笔试题

考试说明：答题时超过 30秒 没有新操作就强制交卷，最多允许切屏 1次 超出该次数会取消继续作答权限。
请同学们注意笔试要求，公正做题。

* 基本信息：

姓名：

学校及专业：

电话：

邮箱：

一、单选题

*1. You have trained word embeddings using a text dataset of m_1 words. You are considering using these word embeddings for a language task, for which you have a separate labeled dataset of m_2 words. Keeping in mind that using word embeddings is a form of transfer learning, under which of these circumstances would you expect the word embeddings to be helpful?

- ☐ $m_1 \ll m_2$
- ☐ $m_2 \ll m_1$

*2. Which of the following do you typically see as you move to deeper layers in a ConvNet?

- ☐ nH and nW increases, while nC decreases
- ☐ nH and nW decreases, while nC also decreases
- ☐ nH and nW increases, while nC also increases
- ☐ nH and nW decrease, while

系统提示

*3. In the word2vec algorithm, you are given a word w and a context word c . How are w and c chosen?

- ☐ w and c are chosen to be nearest neighbors in the embedding space.
- ☐ w is the one word that comes most often in the context of c . 确认
- ☐ w is the sequence of all the words in the sentence before c .
- ☐ w is a sequence of several words immediately before c .

*4. Suppose your input is a 300 by 300 color (RGB) image, and you are not using a convolutional network. If the first hidden layer has 100 neurons, each one fully connected to the input, how many parameters does this hidden layer have (including the bias parameters)?

- ☐ 9,000,001
- ☐ 9,000,100
- ☐ 27,000,001
- ☐ 27,000,100

*5. Suppose you have an input volume of dimension 64x64x16. How many parameters would a single 1x1 convolutional filter have (including the bias)?

- ☐ 2
- ☐ 17
- ☐ 4097
- ☐ 1

*6. You have an input volume that is 63x63x16, and convolve it with 32 filters that are each 7x7, and stride of 1. You want to use a "same" convolution. What is the padding?

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 7

二、多选题

*7. Which of the following statements are true? (Check all that apply) 【多选题】

If the training and test errors are about the same, adding more features will not help improve the results.

If a learning algorithm is suffering from high variance, adding more training examples is likely to improve the model.

If a learning algorithm is suffering from high bias, only adding more training examples may not improve the test error significantly.

A model with more parameters is more likely to over-fitting and typically has high variance.

*8. If your Neural Network model seems to have high bias, what of the following would be promising things to try? (Check all that apply) 【多选题】

- Make the Neural Network deeper.
- Get more test data
- Increase the number of units in each layer.
- Get more training data
- Add regularization

*9. Suppose you have an input volume of dimension . Which of the following statements you agree with? (Assume that “1x1 convolutional layer” below always uses a stride of 1 and no padding.) 【多选题】

- You can use a 1x1 convolutional layer to reduce nC but not nH, nW.
- You can use a pooling layer to reduce nH, nW, but not nC
- You can use a 1x1 convolutional layer to reduce nH, nW, and nC.
- You can use a pooling layer to reduce nH, nW, and nC.

*10. Why do we normalize the input x to Neural Network? 【多选题】

- It makes it easier to visualize the data.
- It makes the parameter initialization faster.
- It makes the cost function faster to optimize.
- Normalization is another word form regularization – it helps to reduce variance.

三、判断题

*11. 当你训练一个视频中描述的对象时，你不需要在训练集中提供，因为算法会自动学习。 系统提示 图片的训练集，然而边框不

- 对
 - 错
- 为防止作弊，只允许在全屏模式下作答，退出全屏算一次切屏。
已切屏1次，超过4次不允许再作答！

*12. 在实现前向传播和反向传播中，我们不需要对权重进行归一化。 确认 间隔值。

- 对
- 错

*13. 给定 n 个数据点，如果其中一半用于训练，另一半用于测试，则训练误差和测试误差之间的差别会随着 n 的增加而减小。

- 对
- 错

*14. 人脸验证只需要将新图片与1个人的面部进行比较，而人脸识别则需要将新图片与K个人的面部进行比较。

- 对
- 错

*15. 回归问题和分类问题都有可能发生过拟合。

- 对
- 错

*16. L2范数会使权重稀疏。

- 对
- 错

*17. 开发和测试集应该来自同一分布

- 对
- 错

*18. 回归和分类都是有监督学习问题。

对
错

*19. 在不同的mini-batch下，不需要显式地进行循环，就可以实现mini-batch梯度下降，从而使算法同时处理所有的数据（矢量化）

对
错

*20. 每个超参数如果设置得不好，都会对训练产生巨大的负面影响，因此所有的超参数都同等重要

对
错

四、问答题

*21. 模型参数（model parameter）和学习算法的超参数（hyperparameter）的区别是什么？

*22. 请大致讲解一下最优中拉格朗日乘子法的思路？KKT是什么？

*23. 看一下下面的这个代码片段：

```
1  ▾ # a.shape = (3,4)
2  ▾ # b.shape = (4,1)
3
4  ▾ for i in range(3):
5  ▾     for j in range(4):
6  ▾         c[i][j] = a[i][j] + b[j]
```

请问要怎么把它们向量化？

系统提示

为防止作弊，只允许在全屏模式下作答，退出全屏算一次切屏。

*24. 深度学习中，有哪些常见的调参

已切屏1次，超过4次不允许再作答！

确认

五、编程题

*25. 给定一个链表，两两交换其中相邻的节点，并返回交换后的链表。
你不能只是单纯的改变节点内部的值，而是需要实际的进行节点交换。
示例：
给定 1->2->3->4, 你应该返回 2->1->4->3.

提交