# 面试手册：

# 技术类：

55 为什么self-attention可以替代seq2seq

在计算attention时主要分为三步，第一步是将query和每个key进行相似度计算得到权重，常用的相似度函数有点积，拼接，感知机等；然后第二步一般是使用一个softmax函数对这些权重进行归一化；最后将权重和相应的键值value进行加权求和得到最后的attention。

可以提取相关的重点信息

56 什么是深度学习中的anchor

目标的期望

57 介绍常见的梯度下降优化方法

**梯度下降(GD)，小批量梯度下降，随机梯度下降(SGD)，Adam**

58 mini-Batch SGD相对于GD有什么优点

mini-batch sgd只随机取n个样本来进行梯度计算，在这n个样本中进行n次迭代，每次使用1个样本更新模型参数，gd是所有样本，所以计算速度比gd快，更能有效地利用信息收敛到局部最优解。

59 什么是感受视野？

感受视野就是卷积核在原图上的一个投影

60 介绍下卷积操作作用

通过卷积运算，是信号特征增强，降低噪声，提取特征

61 CNN权值共享问题

降低参数量，减轻过拟合，赋予对平移的容忍性

62 CNN结构特点

权重，参数共享（共享卷积核，处理高维数据无压力），参数量少，不会受到参数量的约束。

缺点：平移不变性

63 pooling层作用

丢数据，不参与学习，可以增加模型的泛化性能，有几率可以特征平移不变形，降维。

64 深度特征的层次性

65 什么样的数据集不适合深度学习

1：数据集太小，2：数据之间没有局部相关性

66 什么造成梯度消失问题

1：初始权重值过小

2：网络层数过深

3：用了饱和的激活函数

67 Overfifitting怎么解决

1：权重衰减（L1,L2正则化）

2：

68 L1和L2区别

L1：w稀疏化,会达到0

L2：w很难到达0，可以防止过拟合，提升模型的泛化能力

69 TensorFlow计算图

70 BN（批归一化）的作用

**在批次上做归一化，加快训练速度，提高模型精度，防止梯度弥散**

71 什么是梯度消失和爆炸，怎么解决？

梯度消失：权重太小，反向传播时，梯度计算，导数越来越小，神经网络无法被优化，失去学习能力

梯度爆炸：权重太大，反向传播时，梯度计算，导数越来越大，神经网络无法被优化，失去学习能力

解决方法：

1：批量规范化（batchnorm）梯度爆炸：

2：使用Relu Lrelu Prelu Rrelu激活函数。梯度消失

3：梯度裁剪，设置阈值，把梯度限制在一个范围内。梯度爆炸：

4：加残差。梯度消失

5：权重L1，L2正则化。梯度爆炸：

72 RNN循环神经网络理解

记忆，序列。记录同一个神经元在不同时刻的状态

73 图像处理中平滑和锐化操作是什么？

平滑：（中通滤波）降低高频分量

锐化：（高通滤波）增强高频分量，提高了噪声

74 训练过程中模型不收敛，是否说明这个模型无效，致模型不收敛的原因有哪些?

1：数据预处理不当（标注等）

2：数据没做归一化

3：输出层没加激活函数或激活函数导致梯度出现问题

4：网络设计不合理

5：优化器选择不当或者优化器参数不适合

6：权重初始化

75 VGG使用2个3\*3卷积的优势在哪里？

76 Relu比Sigmoid效果好在哪里？

在隐藏层里面，sigmod是饱和函数，容易造成梯度消失，反向传播时除法计算量大，

Relu是非饱和函数，不容易造成梯度消失，正半轴梯度恒等为1，反向传播时计算量小。

77 神经网络中权值共享的理解？

78 对fifine-tuning(微调模型的理解)，为什么要修改最后几层神经网络权值？

79 什么是dropout?

训练时随机让一定神经元停止参与运算

80 dropout具体工作流程？

1：隐藏层随机暂停一些神经元，停止学习

2：然后把输入 x 通过修改后的网络前向传播，然后把得到的损失结果通过修改的网络反向传播。一小批训练样本执行完这个过程后，**在没有被删除的神经元上**按照随机梯度下降法更新对应的参数W，B，然后重复这一过程

81 dropout在神经网络中的应用？

Dropout的网络就是在训练时，通过设置阈值并与某些隐含层节点的权重对比，让特定的权重不工作，加速运算外，防止过拟合。

# 逻辑类：

09 纽约伊沙贝拉时装精品屋，新近从意大利购进了一件女士冬装。这衣服的购入价格再加二成，是该店标出的销售价。出于半个月内未卖出去，女老板又将这个定价减去一成，很快被一位漂亮小姐买走了。女老板获利400元。请问，这件高档女士冬装购入价是多少

售价 = x + x\*20%

def haha(sell1,sell2,n):  
 # x = ((x+ x\*sell2)+((sell1\*x)+ (x\*sell2\*sell1))) - n  
 x = n/(sell1-sell1\*sell2)  
 print(x)  
haha(0.1,0.2,400)

10 一个两位数n,十位是k,个位是5,求（n2 -25）/100等于多少

def cul(k):  
 b = [k,5]  
 n = "0"  
 for i in b:  
 i = str(i)  
 n += n.join(i)  
 n = int(n)  
 print((n\*\*2-25)/100)  
  
cul(6)

11 1元钱1汽水，2空瓶换1汽水，20元最多能喝多少瓶汽水

def soda(x):  
 so = 1  
 bottle = 0.5\*so  
 so = 20\*so + 20\*bottle  
 print(so)  
  
soda(20)

12 如果 A与B是字母，就在答案上写C;只有在5加5等于10的前

提下，才在答案上写D.请问答案结果是A、B、C、D中的哪一个

def guess(a,b):  
 if a.isalpha() and b.isalpha():  
 print("C")  
 elif a == 5 and b==5 and a+b == 10:  
 print("D")  
guess("G","H")

# 项目类：

07 项目交接的流程以及项目交接时发生问题有没有预案？

08 做项目时出现了什么问题

09 尝试过哪些不同显卡训练，速度和批次的问题