

阿里-淘技术bg（凉凉）

一面

阿里一面的时候打电话过来，没接到（囧~），到第二天才接到的。感觉一面问的东西与技术有关的不太多，照例，面试官捞了我的简历之后，还是挺客气的（赞），先介绍了下所在组是做啥的（跟数据分析相关，偏工程的），然后问我是否要继续面（因为我投的算法岗），感觉像是在聊天，问的问题如下：

1. 自我介绍
2. 如何看待，因为偏工程的，而我投的是算法？
3. 有没有学了解spark/hadoop，map-reduce模型？
4. 能不能提前来实习（ps：还没给我发offer呢，为啥要问这个？）
5. 介绍下最近在打的kaggle比赛
6. 给了一个场景题，如何鉴别淘宝刷单记录，要加那些特征？
7. 还有什么要问我（面试官）的？

最后~35mins，结束，说是下周会有人给我打电话，到阿里官网看了下状态，面试中...，应该是过了

二面

到了周二，阿里那边果然来电话了，听声音，感觉是个小哥哥，感觉也基本没问啥的技术问题，像聊天：

1. 自我介绍
2. 简单介绍了下自己的论文
3. 学校的学习模式，如何学习的？
4. 我们部门是偏研发的，你有没有为这个研发做什么准备？
5. 有什么问我的？

最后~30mins，结束，赶紧去看了看官网状态，面试中...，可能是过了吧

三面（2019.8.29）

这个三面真的是让我等得很焦急，到周四下午才来的电话，虽然看着官网状态一直在面试中，但真的是怕莫名其妙被挂了，问到的问题如下：

1. 自我介绍（所有面试共同开篇~）
2. 简单介绍下中的paper
3. 看你研究生期间主要做的是跟视觉相关的，而我们岗很少会涉及到图像场景，这个问题你怎么看？
4. 怎么处理过拟合？
5. xgboost vs. gbdtd?
6. 简单介绍下打的kaggle比赛？
7. 有什么想问我的？

这次问的问题记得比较清楚，刚面完就来写了，最后我问了下后面还有几面，那边小哥哥说我这还是技术出面，后面应该有两轮技术面+一轮hr面，我这真的是有点懵了，这面试流程真是长，后来问了下，可能是我被转组面了，所以这依旧是技术初面。应该是过了吧，看着官网状态还是面试中...

四面（2019.8.31）

这虽然写得四面，但实际上像是二面，因为终于开始手撕代码了，问题如下：

1. 自我介绍
2. 介绍paper
3. 介绍kaggle比赛
4. 问了很多python语言相关问题：
 - 类变量和成员变量区别
 - 函数修饰符有哪些
 - `__call__`方法
 - 自定义python装饰器
5. 手撕代码：连续子串最大累计值

~45mins，面试结束。

蘑菇街（凉凉）

说到蘑菇街，我想吐槽一下，本来我是找人内推算法的，结果给我推到了后台研发，还是Java的，这也太不负责了。还是来回忆下面试吧，两面凉凉，都是视频面试：

一面

一面有点尴尬，我居然忘记面试时间了，以为第二天，然后还是那边面试官给我打的电话，叫我去视频面试，进去之后，问了以下问题，像是大概了解情况的：

1. 自我介绍
2. 还是问我怎么看，我们部门是后台研发，你简历上感觉都是算法相关的。
3. hadoop/spark熟悉吗？
4. 简单问了下简历，介绍了下论文。
5. 来了一道手撕代码，求第k大的数字。

最后~35mins，跟我说下周会有人打电话给我接着面。面完之后，感觉这个面试官好像看不太懂Python，叫我把算法题写个用例跑一下看看结果。

二面

二面面试官感觉就是实实在在搞后台研发了，问的问题都是纯后台的，挂了也不奇怪。问题如下：

1. 自我介绍
2. 线程和进程区别
3. 死锁的必要条件以及如何避免
4. 多线程条件下，如何保证一个全局变量是线程安全的
5. 一致性哈希
6. 来了个算法题，说一棵二叉树，如何找第k小的数字，然后时间复杂度。

最后也是~30mins，确实是感觉没答好吧，尤其是算法题，我说了两种：

1. 直接遍历二叉树，再排序，再返回第k个元素即可
2. 在遍历二叉树的过程中，维护一个大根堆。

第一种算法没啥好说的，比较好理解。然后面试官就追着第二题问我，无奈，对这个大根堆的理解确实不是很深：

1. 先问我这个大根堆的大小和第k小元素是否在堆顶，自己也没理解特别深刻，就不太确定，面试官直接说不是猜谜语（囧~）
2. 又问我时间复杂度，我说 $O(n\log k)$ ，被说不够精确，这是上界。

最后~35mins，直接跟我说，我在数据结构和后台这块比较欠缺，更适合研究。
gg~

依图（凉凉）

感觉依图算法岗还是挺难的，一面二面都是两道手撕代码，题目还不简单，凉凉~。依图面试就比较直接了，直接开始，手撕代码：

一面

1. 自我介绍
2. 手撕代码题：链表重新排序
3. 概率题：一个6面体骰子，一共掷 n 次，设点数之和为 X ，求 $E(X)$

手撕代码题，我说了归并算法，似乎面试官不喜欢归并算法，代码比较复杂，那我说是用冒泡写，我就写了个冒泡，然后面试官接着问，能不能不改变链表节点值，那就改变链表节点next指针喽。然后我就写了个算法，每次扯下原链表中的最小值，用尾插法将该节点插入到一个新的头结点中。

概率题是真的没准备过，这是第一次做概率题，我的思路是根据期望定义去做的，因为 X 范围为：

$$X \in [n, 6n]$$

因此转而求掷 n 次，点数之和为 x 的概率 $P(x, n)$ 。在面试官迭代提示下，用动态规划解决了。

最后面试~60mins，面试官说我把核心代码都写出来了，代码质量也比较高，下周会有人给你打电话。

二面

1. 自我介绍
2. 手撕代码题：两个有序数组，如何求两个数组整体的中位数
3. 概率题：因为那天是七夕，出了一道七夕情景题。说一个广场上一共有50个人，其中有15对情侣（即30个人不是单身），20个人是单身。现在假设一个场景，随机抽一个号码，如果这个号码是单身，就烧死；如果这个号码不是单身，少烧死两个人。求：

(1) 求一次烧死两个人的概率

(2) 如果一直随机抽一个号码，直到一次烧死两个人停止，求抽取号码次数的期望。

手撕代码题，我只是说了一共将两个有序数组合并的方法，然而面试官要我给出一种 $O(n\log n)$ 算法，这个实在是搞不来，想了半天，知道用二分法，算是没做出来。

概率题，第二问也是没答出来，还自作聪明的给自己出了一道题，挖了一个坑。gg~。不过自己也是从来没有准备过概率题，答不出来也正常。

三面

三面去的上海现场面，问了一个手撕NMS题目，又是一个物体检测的，真没做过，写是写出来了，最后花了40分钟，凉凉。

海康威视

一面

一面也是一个小哥哥，很多细节都忘记了，写自己还记得吧：

1. 自我介绍
2. 数据不均衡怎么处理
3. 简单介绍下论文
4. 有什么要问的

聊天氛围还不错，感觉是过了吧。

后来收到二面现场面短信通知，当时已经回到广州了，所以就直接拒了。

美团（凉凉）

首先感谢师兄帮我内推到他们组吧，我连笔试都错过了，僵硬，囧~

一面（2019.8.29）

一面，比较简单，连手撕代码都没有，听说是师兄跟面试官（同事）说我刷了很多算法题，所以手撕代码也省掉了。问的技术问题比我想象中的简单，倒是跟我科普了很多场景题，问的问题如下：

1. 自我介绍
2. 介绍下paper
3. 如何对抗过拟合
4. 如果做特征工程，加特征
5. 如何处理缺失值
6. xgbosot如何看特征importance score，那些特征比较重要，你自己手动加了那些特征
7. 场景题：给我一个手机号，围绕手机号，可加入那些特征
8. 还有什么想问我的

上面的3、4、5、6基本都是在介绍kaggle比赛时候问的。最后跟我讲了很多实际的场景，还跟我说，在实际场景下，反欺诈更重要的是检测出这个欺诈团伙，单笔欺诈的危害并没有那么大。

二面 (2019.9.7)

这一面是总监面，果然没问啥技术细节，直接上问题：

1. 自我介绍
2. 介绍kaggle反欺诈比赛，下一步如何进一步提升
3. 如何评价自己（说了缺点/优点）
4. 给了一张项目首页图，问如何对模块进行迭代，如何做个性化推荐
5. 一些无关技术问题：家在哪儿，工作地点偏好，面了其他公司没，拿offer了没，为啥选美团，职业规划

hr面

当时在做笔试，就如实跟那边hr说了，导致莫名其妙被挂了，也是吃一堑长一智吧。

商汤（凉凉）

一面 (2019.8.30)

emmmm 怎么说呢 反正问的都是我以前没准备过的问题，问的特别细，囧~：

1. 照例自我介绍
2. kaggle machine learning比赛
3. miccai paper
4. 公司实习，各种unet分割网络及其结构特点
5. ISIC 2018比赛
6. BatchNormalization、LayerNormalization、InstanceNorm区别（完全懵逼，没答出来，囧~）现在想想真的是傻逼了，面试官问我给定tensor=(N,C,H,W)，问做BatchNormalization计算出来的mean和std维度是多少，不是经常预处理会计算数据集的mean和std么，维度肯定都是(1,C,1,1)呀（真的给自己蠢哭了）。
7. 假设一个神经网络batch-size=128才能run，但是只有单块gpu，如果让网络跑起来
8. 一个有无穷个元素的有序数组，问如何找到在数组中找到元素k
9. focal loss

二面 (2019.8.30)

二面面试官是接着问的，问的就比较简单了，感觉也比较随意，问题：

1. unet结构和优点，与fcn相比如何
2. 带权无向图，最短路径
3. CT/MRI图像在预处理方面，和自然图片区别
4. 随机森林随机性体现在哪里
5. 优先队列插入一个元素时间复杂度
6. 修改一个堆中元素时间复杂度
7. Bn中的moment参数，SGD中moment参数
8. 还有没想问我的

一面和二面都是接着的，加起来一共1.5h。

云从

一面 (2019.8.30)

面试过程45mins，主要问了两个问题：

1. 自我介绍
2. miccai paper, 非常详细, 还上传了网络结构图
3. 最长不重复子串
4. 蚂蚁上树问题: 一根横杆子上爬满了蚂蚁。已知初始时刻每个蚂蚁的位置、方向和速度。蚂蚁爬行后如果“头碰头”则各自反向继续爬行。求最后一个蚂蚁落下杆子的时间

二面 (2019.9.6)

距离上次刚好一周, 问题如下:

1. 自我介绍
2. kaggle比赛
3. miccai paper
4. 实习
5. SENet
6. triplet loss
7. 实例分割 语义分割 场景分割区别
8. 有8瓶毒药, 其中有一瓶是有毒的, 毒药一个小时后才会有反应, 问至少要几只小老鼠还能把有毒的找出来?
9. 海量数据找最大top k
10. 问了下paper代码都是你写的吗

就这么多, 基本都答出来了。emmmm, 就在这面试半小时里, 错过了两个电话(囧~)

三面hr

hr面, ~15分钟, 问了几个问题:

1. 职业规划
2. 期望薪资
3. 三个关键词形容自己
4. 目前已经有那些offer
5. 你了解到的云从
6. 工作地点
7. paper是一作么

顺丰（凉凉）

一面（20190924）

一面凉经，问的问题如下：

1. 自我介绍
2. paper
3. 解释GAN中的min max
4. 熟悉物体检测吗？RCNN，faster RCNN等
5. don't over fitting

半个小时，直接凉凉。

联影

20190924

一面+HR面

联影是现场面的，一轮技术面+一轮HR面，感觉技术面还是挺简单的，就全程问了问简历，啥都没问；

HR面就一些常规问题的吧：

1. 个人规划
2. 周围人如何评价你的（优点和缺点）
3. 哪里人，家庭情况
4. 个人兴趣爱好
5. 目前有几个offer，期望薪资
6. 期望工作地点

然后在等消息，也不知道啥时候开奖。

电信云计算（凉凉）

一面

感觉这家破公司压根不懂技术，面的算法，全程问研发相关的：

1. 问了一大堆PyTorch相关的，什么合并两个tensor怎么处理...
2. 除了Python，还有其他比较熟悉的语言吗
3. Python继承以及如何子类调用重写函数，如何确定调用那个函数
4. 正定矩阵如何确定的

面试体验极差，感觉面试官压根不懂算法。

华为

2019.10.10,现场面的，直接三面面完，感觉并不是那么简单；

一面

一面感觉比较容易，感觉像聊天，主要还是聊项目，聊简历，最后撕了一道算法题（不用乘法实现两个unsigned int相乘），大概聊了一个小时。

二面

二面相对就难一些，现场很多人也是挂在了二面，刚进候选区，感觉还是人比较多，面完二面出来就感觉人少了不少，所以看样子是二面挂了很多入：

1. 先做了一套卷子，四道题目，有一道手撕代码题目，思路是用二分法求一个数字的平方根（结果取整数），但是限制不能用除法。
2. 复盘笔试
3. 手推bp，SGD/Adam

基本这三个问题，二面大概50分钟；

三面

三面是主管面吧，这个主管比较严肃，笑得比较少：

1. 自我介绍
2. 对华为了解
3. 希望做哪方面工作
4. 高考成绩，哪科考得比较好，哪科比较差
5. 期望薪资
6. 询问我导师情况

将近25分钟，最后握手表示感谢，这次面试结束。

10.25喜提华为算法offer，签约~

深睿

一面（2019.10.22）

深睿一面面了很长时间，面了105分钟，问的问题还是挺广泛的，复盘一下：

1. 介绍一下教育经历
2. 介绍下实习经历，对各种UNet做了比较，能不能说下insight
3. 小样本triplet loss介绍
4. ISIC2018比赛，问的还比较细，一方面问了我们是怎么做的，另外还问我第一名和第二名是怎么做的
5. 数据结构基础，字典
6. 深度学习框架，gpu利用率不高，常见原因可能有哪些
7. 来了一道算法题，说一只青蛙跳台阶，可以一次跳1级，2级或者3级，最多跳50次，跳到100级一共有多少种跳法？
8. 个人基本情况，个人期望，啥时候能过去入职？
9. 检测算法里面，mask rcnn主要的Contribution有哪些？

拿了华为算法offer，二面就直接拒了（第一次这么硬气~，直接不面了）。

腾讯

腾讯是散招，找学长内推了下，听学长说，十几份简历，一个hc。

一面（11.6）

先来了个自我介绍，然后怼了50分钟的机器学习基础知识：

1. Overfitting的表现是啥，怎么解决？
2. L1和L2正则化区别
3. SGD，牛顿法等一些优化算法区别？SGD存在可能陷入局部最优，训练过程震荡问题，如何解决？
4. GBDT和XGBoost区别？工程实现上，GBDT只能串行，但是XGBoost能并行，为啥，具体工程上怎么实现的？
5. GBDT/XGBoost和Adaboost区别？
6. 简单说下随机森林？
7. 你说缓解Overfitting最有效的方法是做特征工程，那么特征工程如何做？
8. Kaggle比赛你用的优化器是SGD，那么理由是什么，为什么不用其他优化器，比如牛顿法？然后接着问，虽然说牛顿法计算量比较大，但是目前已经有解决方案，为啥不用牛顿法？
9. 在带有L2正则项的逻辑回归中，啥是风险损失，啥是结构损失？（我答了，然后我说结构损失类似于XGBoost中的正则项（树高+叶子节点个数）然后接着说比如叶子节点有4个和10个区别是啥（我说10个叶子节点的复杂度更高，产生的loss比较大）？
10. 预剪枝和后剪枝区别？
11. 逻辑回归的损失函数是啥？（交叉熵）为什么可以用交叉熵？（假设服从伯努利分布）怎么判断数据分布是否服从伯努利分布？
12. 最后来了一道算法题，说海量URL，如何找出频次最高的top n？然后又说，如果distinct URL也是海量，如何找最高的top n？