

吐血整理 | 据说这里有一份关于BAT的“宝藏级”面试记录终于可以看了！

原创 守着光的 海边的拾遗者 7月13日

收录于话题

#面试 1814 #面经 190 #互联网 3133 #面试经验及心得记录 19

点击上方“[海边的拾遗者](#)”，选择“星标”公众号

第一时间获取价值内容



平凡之路

朴树 - 猎户星座



这里是归辰的**面经杂货铺**，你想要的都有～

写在前面

作者是研二的学弟，实力强悍的**清北大佬**一枚，研究方向与求职方向高度一致，为**图机器学习/推荐系统**，学弟本人非常有自己的想法，不光优秀还极其自律，不光收割了BAT，还尝试了难度较大的一些外企。话不多说，马上就献上这一份“**珍藏版**”面试记录，内容详实，有一定的门槛，具有很高的含金量。如下仅为部分面经，墙裂欢迎各互联网大佬来撩！（可通过联系小编，具体方法为点击文章顶部左上方“**海边的拾遗者**”进行关注，并在页面右上角选择“**星标**”公众号，然后在页面右下角点击“**联系小编**”即可联系到大佬！）

蚂蚁金服|AILab

一面

自我介绍。

- 介绍实习的工作，问了一些具体内容。
- 了解哪些GNN，GAT和GCN的区别，如果图很稠密的话，对GAT有什么影响。
- 两道easy的算法题，BST和二分。

- 问我会不会一些传统的机器学习方法，比如xgboost, gbdtd。
- 深度学习优化器的区别，Adam和SGD。
- Transformer的结构，positional embedding等。

面试完以后面试官说他挺满意的，介绍了下部门，说主要做图理论算法的。但是我当时比较想做应用业务，所以就让面试官就把我推荐到了他们经常合作的另一个部门了。

二面

到了蚂蚁第二个部门继续面试的流程。

自我介绍。

- 介绍下以前的工作。
- 讲下graph embedding的发展历史，有哪些有代表性的模型，都讲讲。
- 给了两个蚂蚁的场景题，让我设计解决办法，说思路。
- 反问。

总共不到半个小时，聊的感觉挺好的。面试体验很好。

三面

晚上10点多面的。。。。，自我介绍完后简单问了问项目和基础的ML问题，大概10分钟。

阿里妈妈

由于阿里只能最终在系统上选择一个进行投递，我选了蚂蚁，所以阿里妈妈这儿就停止了。

自我介绍。

- 问了python语言的一些属性，list和tuple的区别。dict里key能否是list，如果自己改造的list可不可以作为dict的key。dict除了用hash实现还可以用什么实现。生成器，迭代器，装饰器。
- 代码题：一个随机序列，找第k小的元素。最好的方法应该用quickselect，时间复杂度是 $O(n)$ 。quickselect在之后边面别的公司的时候也被问了好多次，也算因祸得福了。
- 聊了聊项目。
- 问xgboost, gbdtd原理及区别。
- 反问。

字节跳动

一面30min

自我介绍。

- 直接两道算法题：1.判断t2是不是t1的子树（说了暴力算法，面试官说还有更快的。。。被搞蒙了，结果发现他想要的就是暴力递归。最后面试官说我代码有问题，我说没问题，他检查了下说奥是他搞错了。。。>.<紧张） 2.编辑距离，没让写代码，说了下思路。就结束了。

二面50min

自我介绍。

- 只聊了聊几个实习经历的工作，细节问了问。Focal loss公式。
- 一道概率题：给一个木棍，切两刀，问可以组成三角形的概率。有点紧张，想了1分钟说了个错的答案。面试官就提示想想有几个变量？画到二维平面上？就想了到用线性规划，然后就做出来了，1/4。
- 代码题：随机数组找第k大的元素。说了两种，1是用heap，2是quickselect。第二个更快，复杂度是 $O(n)$ ，然后把第二种的code写了下，讲了讲。

三面30min

自我介绍。

- 聊了聊，介绍了3段实习经历。
- 代码题：类似leetcode岛屿连通的那道题。

HR面30min

聊天。

微软|STCA

一面30min

自我介绍。

- 代码题：1.链表排序 2.subarray sum equals k。半个小时写完code后开始聊项目。
- 讲了实习时候做的工作。连带着相关细节数据集的采集，idea怎么做的等等。
- 深度学习的相关知识，过拟合，cnn，rnn，lstm，激活函数，梯度消失梯度爆炸，batch normalization，adam 等等问了问。
- 对面向对象熟悉不，有没有什么经验？

二面30min

自我介绍。

- 代码题：给一个doc和一个摘要，去查找doc里匹配摘要的位置(起始，结束)。能匹配多少就输出多少。比如doc是abc，摘要是ba，那么b和a都是在doc里有所匹配的。如果可以匹配多个，则返回最长匹配。

一开始没理解题意，问师姐怎么划分摘要里的句子，用句号？后来明白了不需要划分，就硬匹配就行。面试官提示我建立一个倒排表，记录doc里每个字符出现的位置，才终于明白了。coding了一下讲了讲，算法就过了。面试官说ok，coding没问题算法就这样吧。

- 介绍实习经历。简单聊了聊实习做的东西，大概十分钟。整个面试体验挺好的。

美团

一面40min

自我介绍。

- 上来先介绍实习经历，项目。做法，细节，创新点都有问。从我做的东西出发，问了一些深度学习的问题，比如学习率，优化器，BN，LN，梯度消失爆炸，过拟合，正则化的方法等。如果训练的时候不同batch之间loss差别很大可能是什么原因？当embedding size特别大的时候怎么办？
- 后边做了一道非常简单的题。给一堆学生，按成绩排序。
- 问了python的语言特性，*args, **kwargs, yield, 生成器，迭代器，python的垃圾回收机制。知不知道python多线程。python构造函数是啥？

二面40min

面试官看起来略微严肃，问了很多问题，问的很细。

自我介绍。

- 聊项目，基础。
- 还有一些发散性的问题。比如怎么由用户的外卖信息对用户作其他方面的推荐（比如旅游）。
- 代码题：求topk小的数。

Amazon|SDE

一面

自我介绍。

- 介绍项目。
- 代码题：1.minstack 2.system design的题目，是最基础的LRU。但是当时没做过，靠着面试官提示，半个小时最后终于把思路说对了，没时间写代码了。

二面

介绍项目。

- 代码题：给一个list，里边有一些time，比如23:59，14:00，让你返回里边最小的时间间隔。在面试官的指引下，最后把时间复杂度降到了 $O(n)$ ，因为实际上可能的time只有1440个。

这题回头看挺简单的，唉，但当时面试的时候自己好菜，就得靠面试官提示才到了最全，最优解。最后面试官说我potential还不错，他后边还要面很多人。等通知吧。

因为是SDE岗，感觉面试官倾向于聊计算机基础，比如网络，数据库之类的。。。

腾讯|WXG

一面

自我介绍。

- 写一个堆（好久不写了，都忘了堆了。。。写了40min才写完，发的那个在线coding的地方写代码会有重影眼花了都。。。写的有一点小问题，被check了一下，不过应该还ok）。
- 问深度学习基础。上来问我为什么要引入激活函数。我说是为了非线性。他说错，根本原因是为了scale输出值防止梯度爆炸。我？？？然后问我sigmoid是非线性的吗，我说是啊非线性的。他说不，sigmoid, tanh都是线性的，relu才是非线性的。我蒙蔽了，他又跟我讲一开始引入sigmoid为了scale值，防止梯度爆炸/消失，balaba。。。我无语了只能同意他了。。。
- 问batch size有什么用？大的batch有什么优点？
- 问基于字和基于词的nlp方法有试过吗，哪个好？
- BTE听说过吗，我以为他问的BCE，又仔细问了一遍，确定没听错是BTE。不知道，后来查了下也没查出来是啥。莫非想说bert？但他是一个一个字母读的，我也没听到啊。。。。
- 问nlp里decoder出来unknown词怎么办。我说decoder的softmax那里应该把unknown mask掉的，就不会输出这个了。如果非要输出，就再训练一个模型学一下这个unknown应该是啥。他说ok吧。。。

- 然后时间差不多快到了，就问我有啥问他的。我说有做推荐搜索的吗，对话这个方向可能不是很match，我不了解。他说也有做推荐搜索相关的，简单介绍了下。然后就结束了。。。

二面

刚开始自我介绍我说想做推荐搜索的，和对话系统不太匹配，面试官就把简历给HR推到比较合适的组了。

自我介绍。

- 问了一个python的函数默认参数的问题。答错了，面试官让我打开命令行自己跑一下，然后给出解释，并且改成这个函数应该想达到的效果。然后连着调用两次 `append_list('one'), append_list('one')`，问两次返回啥。应该 `['one'], ['one', 'one']`。函数如下：

```
1 def append_list(new_item, a_list = []):
2     a_list.append(new_item)
3     return a_list
```

但是其实这个函数的功能是想两次都返回 `['one']`，所以我改成了：

```
1 def append_list(*args):
2     if len(args) == 1:
3         return [args[0]]
4     args[1].append(args[0])
5     return args[1]
```

- 代码题：找到两个list相交的点。
- 接下来半个小时主要针对我之前的工作聊了聊，包括一些细节。
- 最后问了一个发散的问题。对于一个搜索query，如果这个query很长，里边很多没用的词，那么怎么提高搜索的准确度呢？我主要从模型上讲了讲，可以build一个模型来确定query里每个词的重要度，用cbow/self-attention之类的模型。
- 最后面试官介绍了下他们团队，感觉不错，接下来等通知后续安排。

三面

自我介绍。

- 介绍项目。
- GNN发展史，GNN为什么可以work，GNN的数学原理。
- self-attention的深层次原理是什么（不懂。。。），为什么work？

- 代码题：一个单调栈的题，比较简单，他那边测了下test过了，然后就结束了。

HR面

自我介绍。

- 问我你的缺点是什么？结果被追着问了好久，第一次有hr面要翻车的感觉。

不过好在hr最后说今天是腾讯提前批的最后一天，应该一会儿你就能看到状态更新，下周会发offer，看来是通过了。晚上11点多收到了短信说恭喜通过了所有面试，具体offer沟通下周进行。

快手

一面40min

自我介绍。

- 机器学习/深度学习基础，问了好多，有一些不会。
- 聊了聊项目，面试官对我做的方向不是太了解。
- 代码题：带重复元素的sorted array里寻找最左端的target。

二面1h

一面通过了马上就开始二面，感觉这轮的面试官特别聪明有活力。感觉这轮主要考coding。

自我介绍。

- 问了一些python的基础，list, tuple, dict, generator, derocator之类的。
- 代码题：1.zigzag打印二叉树 2.给两个等长字符串，输出两个字符串之间的个数 3.给第三个字符串，求两字符串中间不包含第三个字符串的个数（leetcode上周周赛最后一题。。。太难了）不会做。就把第三个字符串简化为字符，和面试官讨论了一下。

Apple|Maps

一面20min

自我介绍。

- 简单介绍了下简历项目，说接下来还有一轮过项目一轮coding。等通知但是后边没消息了，应该是今年疫情的缘故，HC取消了吧。

总结

整个春招过程可以说经历地非常快了，体验也还可以，有失也有得，后面的路走一步看一步了。小编觉得优秀的人本质都是非常相似的，向优秀的人看齐才是正确的姿势，但是更优秀的人更努力的例子不值得我们更加努力学习吗。最后祝大家面试顺利，一起加油！欢迎大家来交流～

▼ 更多精彩推荐，请关注我们 ▼



写留言

来都来了，喜欢的话就请[分享](#)、[点赞](#)、[在看](#)三连再走吧～～～

