

2024 年初到现在

大模型岗位面试全纪录，分享经历与感受

1、字节 AML (offer)：

字节一如既往的注重 coding，基本每一面两道题，自己能拿到 offer 也算是运气吧。一面的面试体验很差，可以说是毫无面试经验的面试官。

2、边塞科技(拒)：

这家专门做 RLHF，微调的。清华背景，创始人有 openai 的经验。聊了两面感觉在 RLHF 上积累特别多，学到了很多。楼主 RL 是个票友级别就无了。

3、阿里达摩院(新达摩) (offer)：

一面考了 leetcode，浅问了大模型技术。第二面是团队大老板，海归，之前在北美连续创业成功，汇报行癫。聊了很多高层的对于大模型技术上的愿景和理解，大模型未来市场方向之类的，我其实听不特别懂，不过学到了很多。他家是纯研究方向，算力听意思挺 nb 的。主要做大语言模型和多模态大模型这两方面的。最后说我一面 leetcode 不太行，补了一面 coding，然后收到了 offer。

4、月之暗面(moonshot) (拒)：

这家感觉面试问的广度和深度是国内厂面的里边最相关最专业的。一面是个大佬，之前 Kaiming 大神一些高引论文的共同作者。一开始问了包括一些基础的大模型训练的内容，transformer 架构。然后开始问一些框架端的内容，分布式训练的切割和 device 之间交流，后面 follow-up 的问题问了不同 device 之间 communication 怎么做，底层 communication 算法有啥，我答案里有 ring-reduce，然后再深入问了 reduce 底层如何实现，我就不会了。问了会不会用 cuda 写东西，我说了解很浅。最后给了比较难的 leetcode，题目不记得了，没写出来，挂。看了下他家主页，感觉创世人和员工都很牛。这家虽然只有一面感觉在国内应该是背景很 top 的一家了。很可惜没有机会多聊一些。

5、零一万物 (拒)：

一面是之前阿里的推荐系统那边的大佬，问了 leetcode 和一些推荐系统的问题，感觉答得还可以。二面也是个大

佬有搜广推，NLP 相关很多经验，考了个需要动点脑子的 leetcode,没答上来，挂了。没有走到很靠后，好奇他家后边面试会不会有大模型的大佬来面。

6、好未来(offer)：

他家感觉团队整体背景在众多公司中相对一般，一面问了一个简单的 leetcode 没答上来很尴尬，模型轮可能比较好最后还是给 offer 了。问题也比较中规中矩，包裹没聊。做的项目是数学相关的 GPT 模型，个人也不是很了解这个方向。

7、IDEA 研究院 (拒)：

这家确实啥也不记得了，记得的就是一面就被拒了，很尴尬。

8、科大讯飞 (拒)：

HR 先聊的，觉得我不可能来合肥就拒了。

9、百度文心(offer)：

这边面了好多轮。一面是数据组的人，问了很多关于大模型数据处理的内容，感觉答得一般，表达了想去模型组，然后面试官应该是和 HR 交流了，第二组开始就是模型组的人来面试。有没有 leetcode 不记得了，面试内容也比较中规中矩，transformer，大模型训练这种内容。后边和 leader 还有 HR 聊得也比较通畅。和初创不一样他家组分得挺细的，模型，框架，数据啥的都是分开的。(包裹底薪 4 万多，几个月忘了)

AI 大模型学习路线大纲+系统学习课程+必备工具，需要的小伙伴扫描下方二维码，我会给你发的

哈~ WX 号：qt02746



10、百川智能(拒)：

猎头推的，安排了面试，第一面过了之后，后面二面我记错了时间，晚了很久(在外边开车)。然后被 HR 拒掉了，

不过理由说的是要 3 年以上经验的，不知道是不是给我点面子。

11、商汤 (没消息了, 被拒):

一共两面，一面记得是 leetcode+写 MHA，感觉面得不错。二面好像答得不太好，具体问的啥忘掉了。。。后来没消息了，被拒。

12、小红书(简历拒)：猎头说要 phd。

13、智元机器人(Agibot) (拒)：

稚晖君那个。当时面的很早，他们好像还在一边组建团队一边搞东西，一面就是稚晖君本人，问了一些简历的经历，然后问了 transformer 相关的比较常见的问题比如 MHA，复杂度，BERT/GPT 这种。然后后边就没有收到回复，应该是挂了。这家是我开始面试的第一家，所以印象还比较深刻。和大佬交流感觉也挺顺畅，很看好他家的以及具身智能这个方向。给稚晖君大佬点赞！

14、面壁科技/面壁智能 (offer)：

记得一共两面，技术上问了大模型训练，transformer 相关的内容和过往经历，第二面好像就聊 offer 了如果没有记错。团队很年轻，大部分清华背景。当时感觉是不是有点太年轻了虽然清华背景这样子。印象里他们有说 funding 还是很足的。PS:包是挺大的。

15、光年之外 (简历拒)：

三月份猎头帮忙推荐的，简历拒。

16、北京智源人工智能研究院 (消失)：

一共 3-4 面，面试官是不同 team 的人，有 evaluation，训练等团队。面试时总体感觉他们想招偏 evaluation 这边的人。面到后面记得有一次面试官鸽了约好的面试，然后后面说从新安排时间，但是猎头或者 HR 就没有再给后续消息了，应该是默拒了。

17、360(简历拒)：当时猎头说要高 P，P8+

18、Minimax(口头 offer):

这家面试轮次挺多，具体记不清但是起码 4+ 轮。前几轮都是考 leetcode 为主，问了很少很浅的大模型的内容，记得是其中一面是让手写 MHA。最后一轮和主管聊完之后被分配到了框架组的 offer，和猎头问了一下算法组是不是

考虑，回答是不，就没再继续接触。PS:他家的包还是挺大的（总包说肯定到得了 100+），在几家初创里都算大的，但是面试时公司的前景，业务技术方向感觉面试官没有沟通很多。主管面也是 20 分钟就匆匆结束了，没有交流很多信息。

19、昆仑万维(offer)：

面试好像问了之前训练大模型的内容，比如模型参数大小，训练一些细节之类的，印象里没有 leetcode。面试官记得很乐呵的一个人。当时拿到的包裹相对低一些(65 总包)比其他公司。不过后边接触到猎头都说他家给的还是很可以的，所以不确定是不是自己记错了或者交流有误。

20、云从科技(拒)：

两轮面试被拒掉了，印象很深的是第二轮面试超级长，设定 1 个小时面试聊了 2 个多小时。面试官问了很多之前的经验，大模型训练的细节之类的，模型端，框架端，之前做的研究都聊到了。感觉聊的是挺开心的，不过后边就被拒了。

21、阿里夸克(offer)：

4 轮面试，2 轮带 leetcode，然后浅浅问了相关经验，第三轮主要问了大模型的内容，比较标准的问题(transformer，训练，分布式，如何处理训练一些问题比如 loss spike 啥的)。第四轮是 cross 面，推荐组的一个大佬，感觉是象征性面试，因为之前做过搜广推问了相关问题，问了 probability 和脑筋急转弯，最后问一些 OS 的问题我不会。几天后收到 offer，和 HR 接触时感觉不太礼貌，交流有些摩擦。包裹差不多 70-80 之间的算上签字费)。

22、衔远(offer)：

3 面左右，问的也比较中规中矩。最后一面是和周伯文老师聊。他家最后给了框架的岗位，但是我相对想做模型相关。包裹挺大的。

23、潞晨科技(Colossal-AI)(拒)：

他家是做框架为主的，一面问了很多框架内容，各种模型切分方式，问的比较深，好像也问了 flash-attention 的内容，感觉答得还凑合。后来我表达了想做算用端比较多，感觉可能不做基座?然后面挂了。他家在圈子里感觉框架端名声挺大的，记得公司内部之前有人发他家新闻来着。

24、蚂蚁(offer)：

徐鹏老师团队。一面记不清了，二面是徐鹏老师本人。聊了之前很多经历，大模型，搜广推，本科实习时的前后端。

印象深一点是徐鹏老师表达了团队人希望模型研究和工程都要做，research 和 engineering 不分很细，这个和我的

观点很像。接下来一轮 HR 面就发 offer 了，第一年算上签字 70-80 这样子。

25、腾讯(简历拒)：

猎头说要 phd。

感悟：

- 1、大模型这方向真的卷，面试时好多新模型，新 paper 疯狂出，东西出的比我读的快。
- 2、Research 岗位对工程也有要求，工程端也需要了解模型。
- 3、感觉比较硬核的岗位，尤其初创公司都是对好几个点都有要求的（应用，模型，框架，底层后端,硬件）。
- 4、楼主目前是:模型>框架>底层>其他。下一步想不放下模型的前提下，发展底层这边的能力(kernel 等)
- 5、目前市场还看不太清，楼主对市场之前不太关注，也不感兴趣，技术宅。6、RLHF 很有前景的方向，强化学习如果经历够的话也要深挖一下。
- 7、大模型包总体给的相对高一些(同 P)。
- 8、目前大多数公司还是集中在语言模型，偏研究的方向会有多模态的预训练这意思。