

# 面经总结

8.9 正式决定 quit 重申，开始找实习。首先建议尽量找内推，系统流程很慢，有 coding 面还要刷题很麻烦。先介绍下几家大厂 NLP 和语音岗的找实习方法和条件对比，然后讲下我在 MSRA 和阿里达摩院的面试经历，最后总结一些面试经验。

## 公司介绍

### 1. MSRA

Website: <https://www.microsoft.com/en-us/research/lab/microsoft-research-asia/groups/>

(备注: MSRA 官网没有语音组, 语音方向都在多媒体, 机器学习和自然语言计算组下去找)

Base: 大部分岗位在北京, 少量在上海;

Salary: 5k/月 + 2k/月房补;

Employment: mentor 有 hc 会在 AI 求职, MLNLP 英才汇等公众号发文招实习, 也可以上官网找中意的 mentor 点对点联系, 流程类似申请学校找导师。

### 2. Alibaba DAMO

Website: <https://talent.alibaba.com/campus/position-list?campusType=talentPlan> (校招主页)

<https://damo.alibaba.com/labs/> (达摩院实验室官网主页)

Base: 大部分岗位在北京, 少量在杭州;

Salary: 400/天 + 50/天餐补 + 2k/月房补;

Employment: mentor 有 hc 会在 AI 求职, MLNLP 英才汇等公众号发文招实习, 无 mentor 个人主页, 不能点对点联系, 但同一个组不同方向都是单独招生的。

### 3. Tencent AI Lab

Website: <https://ai.tencent.com> (官网主页)

Base: 大部分岗位在深圳, 少量在北京;

Salary: 不详, 貌似比 MSRA 较强, 比达摩院较差;

**Employment:** hc 很紧张，公众号上很少见招生信息，见过个别 mentor 在知乎上公开招生，组别规模明显比 MSRA 和达摩院要小，只有 NLP 组有主页可以看到 intern 信息，其他组申请只能发官方邮件（极不推荐）。

其他申请页面：

字节：

[https://jobs.bytedance.com/referral/pc/position/share/?keywords=&category=&location=&project=&type=&job\\_hot\\_flag=&current=1&limit=10&functionCategory=&token=MjsxNTc2NzUwNzU0MDkxOzY2ODc5NDcxMjgxMjUyNzc3MDA7Njc3MjA4OTQxNjc5NTA2NDU4OQ](https://jobs.bytedance.com/referral/pc/position/share/?keywords=&category=&location=&project=&type=&job_hot_flag=&current=1&limit=10&functionCategory=&token=MjsxNTc2NzUwNzU0MDkxOzY2ODc5NDcxMjgxMjUyNzc3MDA7Njc3MjA4OTQxNjc5NTA2NDU4OQ)

百度：<https://talent.baidu.com/jobs/list?recruitType=INTERN>

## 面试流程

这段时间面试过的有 MSRA 的自然语言计算组（NLC），机器学习组（量化金融），腾讯 AI Lab 机器学习组，阿里达摩院语言技术 Lab 的对话系统组和多模态翻译组，还有生物信息方向一个组，还有百度 NLP 组，商汤 NLP 组，还给字节 HR 刷了个 KPI 给了个简历挂。大部分组都是练手，只过了一面就没下文了（这里建议大家不要海投练手，你们实习和秋招离得比较近，这么做很败人品，我是下次秋招距现在几年后了不存在这个问题），走了完整流程的只有阿里达摩院的对话系统组和 MSRA 的 NLC 组。

### 1. MSRA NLC Group（在 Team 进行，两轮技术面+一轮主管面）

一面：决定润第二天就投的，所以准备很仓促，建议 MSRA 这一 level 的公司不要拿来练手，不然白浪费机会。

项目介绍（20min），聊得很不错，主要做预训练语言模型，方向挺 match。第一次没准备 PPT，所以细节没聊太多，主要问了些我都做了哪些工作，想出哪些 idea，和以前 work 的对比等等。

Coding：会发一个网页链接可以公共编辑。第一道是写语言模型计算 PPL 表达式，写出公式表示即可；第二道是计算网格中从原点到特定点的最短路径长度，leetcode 1091 题，需要广度优先遍历，难度中等，由于事先没准备 coding，完全懵逼，我解释说本科自动化没有学过数据结构，这里温馨提示没 CS 背景的别和 CS 硬扯关系给自己脸上贴金，

不然做不出数据结构不好解释，我就多事的写了个本科专业是人工智能与自动化，然后瞎扯我们专业是偏电气电拖自控的，没学过数据结构。然后启用微软复活甲机制，第二天又补了一轮 coding 面，这次是数组中的 0 移到末尾，leetcode 283 题，虽然简单但还是手生了，虽然写出来但还是有点超时了感觉。

二面：一面过了三周，我以为被刷了还是被忘了，结果又被想起来了。这里感觉是因为距离下一季实习生入职还有一个月，可能为了等更多人竞争，所以被吊了很久。

项目介绍，这次做了 PPT，所以讲得比较详细，主要关注点还是 Motivation，和以前的方法对比，然后 Bayes 方法 work 的原理，最后谈了一下对语言模型发展的一些看法。

技术考察和 Coding：先让我说了下 Transformer 的结构，就所有子模块及其功能，包括位置编码的实现，为什么用 LayerNorm，Multi-head 的作用，Attention 计算后 Softmax 的作用，FFN 映射到高维的作用，有没有非线性单元在哪。这是非常高频的题，建议仔细阅读 Attention Is All You Need；然后考了个 coding，就是伪代码实现 Word Tokenization 和 Embedding 的过程，不用运行，写出思路即可，这种题不一定代码要写的多准确，主要是思路清晰，把思路写在注释里，标明步骤，即使代码有错也无伤大雅。

主管面：这一轮是电话面，主要就谈下简历，然后自己 PhD 的学术目标，为啥转 NLP，都有过哪些了解，还问到有没有用过 multi-GPU，说他们都是几十张卡并行的，让我适应下用多卡。最后确定了入职时间和实习时长，然后说要啥准备材料😭😭😭，听到导师同意函就拔凉了。被你们下午一吓晚上没绷住去自爆了，后面就无了。

点评：感觉没录我并不是公司规定，毕竟 MSRA 作为中国业界科研院所的 AI 四大天花板之一（MSRA，华为诺亚，阿里达摩院，腾讯 AI Lab —— 博阳评），还在北京，清北每年的 NLP 实习生就不少，像我这种又是换方向又是非清北的，背景吸引力不够，很难让人家破例；如果真是因为死规定才把我刷了，那我反而觉得 MSRA 在走下坡路了。

## 2. Alibaba DAMO NTL（在 AliMeeting 进行，三轮技术面+一轮 HR 面）

一面：发了简历后过了一周才安排面试，有时候得适当提醒下面试推进。

项目介绍（30min），这次做了 PPT，讲得很有条理，主要关注点是 Motivation，以及我的 contributions，还是和 xdm 上次说的，用通俗的语言讲出来，让他快速能理解，不然他就会觉得你没理解，比如我做的语音，他就很关心类似的方法在 NLP 上有没有相

关研究，为什么不适用，语音和文本 gap 这些，很多靠临场随机应变。反正他也不一定懂，能自圆其说就行，说话要很有自信。

技术面：考了一堆机器学习基础知识，运气很好这面试官也是个 PRML 爱好者，所以这次面试完全逮着了。

1. L1 和 L2 正则化怎么实现（损失函数后加上参数的绝对值求和（L1）或平方和（L2））；怎么起作用（增加模型的稀疏性）；哪个更稀疏（L1，L1 的收敛区域是个方形，L2 更平滑）；L2 怎么推导的（最大后验估计，参数假设为 0-1 高斯分布）；L1 怎么推导的（这题不会了，随便扯了个均匀分布，其实是需要用共轭来计算，具体形式他没算过）；这题对应 PRML 3.1 节线性回归一个图，专门有对比，他也看过书想问的就是那个图，感觉能答出稀疏性就已经可以算过了，后面那个 L1 推导应该就是随口一说没答上来不要紧。
2. 判别式模型和生成式模型的区别（判别式直接对后验概率建模，生成式对联合概率分布建模，也就是分别对似然项和先验项建模）；说一些我了解的生成式模型（刚好看过 diffusion model，把 VAE，GAN 也都说了一下）。
3. SGD + 动量法和 Adam 哪个更好（这个我实验里两个都差不多）各自的原理（SGD+动量法：动量法采用指数加权移动平均，认为当前梯度只与之前多少步有关，然后使用过去梯度的均值代替当前梯度，减少震荡加快收敛；Adam：自适应调节学习率，指数加权移动平均估计梯度的动量，通过对梯度的累加来调节学习率，只说了这些具体推导过程太难了）。
4. LSTM 和 Transformer 如何进行序列建模（LSTM 是 Recurrence 结构，不断把输出递到输入，Transformer 是前馈结构无法直接建模，因此在 word embedding 上加了位置编码，对于生成式任务需要在 attention 矩阵上加掩码）。

虽然有些答得不完整，但整体感觉挺不错，面试官也说基础还可以，问了入职时间后，说 OK 了，去系统里走下流程吧，然后走到现在了。

二面：一面过了一周多，中秋前安排了二面。

项目介绍 PPT，这个面试官问了很多实验细节，还有我写文章的心得，对于我干了多少活挺感兴趣，倾向独立自主科研还是合作式，还是唯一一个对我本科智能车经历感兴趣的，问我在团队里怎么协同，发生分歧会怎么处理，感觉提前把主管问的都问了。问我

对对话系统了解我说不咋了解有点伤。就我是想转 NLP，但之前没接触过所以具体分支不怎么了解，觉得模型相近的话应该可以尽快上手。

技术面：

1. GPT 和 BERT 的区别（GPT 生成模型，以 Transformer 解码器为基础，适合语言模型等任务；BERT 是双向结构，以 Transformer 编码器为基础，适合文本摘要等任务）
2. Transformer 的结构，这个问得更细，除了上面提到的关键点，还问了我代码里怎么实现 multi-head 的（就是映射成多个 head，每个 head 单独进行 attention 运算，最后 concat 一下再映射回去）LayerNorm 和 BatchNorm 的区别（BatchNorm 是对不同 batch 同一个 channel 上来 Norm，LayerNorm 在 NLP 任务上是在一个 token 的所有 embedding 上做 Norm）。
3. 问我平时是否关注其他模型，我说了 PRML 里的基础，还有 DeepLearning 的热门模型，比如 VAE，GAN，GNN，还有最近的 diffusion model 等等。这里防了一手，我感觉没到聊天环节，应该还是在考技术，所以没太吹的天花乱坠，说的都是我比较熟悉的，果然，接着就让我讲下 GNN 和 diffusion model。所以面试还是尽量低调些，这题老实回答就是送分题，乱吹就是送命题。

也是当场就说过了，下一轮主管面去安排一下后面联系。

三面（主管面）：还是先 PPT 介绍项目，重点问了 motivation，和 NLP 任务的联系，怎么实现的，这里说了半天，我没 get 他的意思，他想问我们方法效果提升有没有例子，我给他说是语音识别上准确率提升，他就说想看具体的例子，后来我说这种方法在单个句子上看没有意义，关注的是统计充分性测试结果，因为输入的扰动太大了。这里感觉是故意刁难这么问的。然后就是具体实验细节，问了些 Dropout 的实现原理，贝叶斯优化的了解等等，没考技术，问了下入职时间，职业规划等等，就过了。

点评：不考 coding 相对还是轻松一些，就是太慢了，拖了一个多月了，从我期刊大修投递申请，到最终版见刊了还没走完流程。不过也只有第一轮考得比较难，剩下两轮都是比较轻松聊天式进行的，剩下 HR 面又拖了两周，最近实习转正太多排不开，HR 面基本不挂人，就问些家庭情况，思考规划之类的，女生据说会被问有没有男朋友，只要不胡说八道都会给过。

## 面经分享

1. 时间安排上，入职前两个月前就可以投了，最好找熟人内推，商汤都是明写了有内推直达 HR 面。
2. 准备方面，一定要做 PPT，适当吹牛，别吹过头就行，有无 PPT 简直两种体验，PPT 讲得多了考技术时间就少了；然后最好找一些没打算去的练练手，我除了 MSRA NLC，达摩院对话系统和腾讯 AI Lab，其他都是用来练手的。很重要的一点是 PPT 介绍和修改，就是一个自监督训练的过程，和面试官交流后你才能知道他们更关注哪些点，把一些没用的信息去掉，突出重点，比如我的 Motivation 还有 Related work 就修改了很多遍。我最初 Bayesian LM 这部分就是用 ICASSP 那个 PPT 介绍的，但很多技术细节根本没必要讲，把文章 Abstract 能讲明白就已经够了。后面对时间把控也越来越好了。
3. 跟 xdm 之前说过的，平时多开阔视野，多联系，多思考，很多问题面试官也没答案，就是考的临场应变和自圆其说，说话要自信，气势上不能输；然后多用通俗化的语言并非书面表达那样跟别人解释自己的 motivation，因为文章里的表达都是高度凝练建立在大量参考文献基础上的，虽然更系统专业但不适合短时间内让别人去理解。这里建议我们 seminar 可以组织类似的，给别人介绍自己 work 的 motivation，我们就像面试官那样去问。如果平时一直和内行的比如老板去聊的话，你会忽视掉基础细节，多和外行聊一下，会发现很多没注意的问题，如果都能解释的通，理解绝对要上一个层次。
4. 以后想到什么再补充。

**最终去向：华为（暂定），感谢鸿儒老铁送的 HR 直达！**