

阿里

一面

没打招呼突击电话面试（30min）

- 地点ok? 时间ok? (base杭州, 暑期, 全部ok ok ok) 简单自我介绍一下, 项目经历选一个说说?
- 项目中对于没有标签的数据, 怎么处理, 还有什么别的方法? (我用的远程监督, 别的还有one-shot zero-shot label embedding)
- 假设一个情景题, 微博创立初期, 短文本, 10w数据集, 无监督, 分类怎么做 (我答的用label embedding做或者用embedding+聚类kmeans)
- embedding怎么实现? (fasttext or word2vec, 不能用BERT, 数据量大不合适) 补充: 你说的是词向量, 句子向量怎么做呢 (+average pooling最常用, 还有别的方法没列举)
- 除了你说的方法, 还有别的方法可以吗? U-learning听说过吗? (没听清? 啥learing? 这个真不知道, 只听说过Q-learning)
- kmeans的k如何选择, 这么做效果好吗? 为什么? (k不知道, 具体情况具体分析吧 (编不出来了), 效果不好)
- 那如果不用聚类, 别的方法怎么做, 提示: 抽取关键字 (关键字抽取可以用词与句子做相似度匹配)
- 按你说的不具有通用性, 那如果是长文本, 又怎么做, 有什么区别 (长文本用词频统计+停用词过滤)
- 你没回答出我想要的答案, 因为一个微博可能属于多个类别 (多类别无监督分类)
- 快排知道吗? 稳定不? (不稳定, 因为会交换顺序)
- bagging、boosting知道吗? (知道, 属于ensemble, 叽里哇啦叽里哇啦)
- 用啥语言, 用啥框架 (python, pytorch)
- 反问: 实习时间?
- 反问: 表现如何?
- 个人体会: 面试官人超好, 很客气。感觉工程能力还有待提高, 很多东西还是停留在理论上。kmeans自己挖的坑, 无语了。。。后面问了大佬, 这个情景题可能是LDA之类的主题模型, 或者是手动标记数据, 再训练。或者是用迁移学习。

二面

约好的电话面试（30min）

- 自我介绍，问论文（简单介绍一下）
- 问比赛，效果怎么样，名次之类的（问啥答啥）
- 二叉树有了解吗，遍历有哪些（前中后层次）项目用过哪些数据结构（好像没有，都是用python做的）
- 动态规划了解吗，解释一下，填表是怎么做的（吧啦吧啦）
- 操作系统了解吗？（本科学过，现在忘了。。）
- Linux会用吗（会，一般用来跑实验和配环境）Linux debug的时候怎么做的（我用的nohup，输出log日志来debug）
- 计算机网络学过吗？（本科学过，现在忘了。。）那学过哪些课（ML DeepL NLP）DeepL的课程作业有吗（没）
- BERT模型说一下，历史、演变说一下（RNN、LSTM、ELMo、Attention、Transformer、GPT、BERT、Auto-Encoder、Auto-Regression、Positional Encoder）
- 情感分析怎么做的说一下，对你后面有没有什么启发（根据项目说的，样本不均衡、评价指标等等）
- 别的语言会吗，java, c（本科学过，熟练度不高，现在捡起来很快）
- 有没有关注什么学术平台，或者线下的学术活动（知乎、群、机器之心、自己组织线下交流活动）
- 反问：表现如何
- 反问：钉钉这边的业务是什么
- 个人体会：面试官很客气，像聊天一样，很顺畅。

携程

一面

- 自我介绍，实习时间，22年可以嘛（22年看情况）
- 分词知道嘛jieba说一说（知道，维特比算法巴拉巴拉）
- 分词种新词发现怎么做？（几里哇啦扯了一堆）你说的好像和我不是一个东西，新词发现有哪些常用方法（啊。。。不知道）
- 维特比算法中的DAG知道吗，怎么构建的（知道，扯了一堆）
- 所以DAG怎么构建你没有说（构建。。？不是根据统计信息得到的嘛？不知道）

- 看你熟悉维特比算法, HMM知道不 (知道的... (西麻哒!!))
- $\log P(O|A,B,\pi)$ 怎么变成只有 $AB\pi$ 的式子 (啊这。。。扯了一堆, 三个问题, 参数估计, d-separated, 动态规划 (完全扯错了方向))
- 感觉你说的和我不在一个频道上 (哭, 真忘了, 没想到会问这么细)
- 看你论文用了BERT, 说一下BERT和transformer的positional embedding有啥区别 (多了一个矩阵, 多了一个dropout几里哇啦)
- 这个positional embedding怎么生成的 (cos, sin生成的, 相当于一个look up tabel)
- positional embedding公式写一下吧 (啊, 这个没背啊)
- sigmoid当数据太大的时候不敏感怎么办 (normalization)
- 最后做个笔试, 用np实现一个LSTM, 初始化权0.5 (啊, 能不能写伪代码?) 不能写伪代码, 要跑出来。 (啊这。。。我好多np函数用法没背下来怎么办) 你让我有点惊讶, 居然没记下来这些常用方法 (啊这个sigmoid怎么调用着?) 你自己写 (啊好的好的, 还好记得公式)
- 能运行下不 (可以, 然后炸了) 你这bug在哪你知道不 (应该是矩阵搞炸了) 是的, 怎么改 (矩阵操作函数名忘了) 好吧我提醒一下, 把math.exp改成np.exp (啊好的好的) 你这又有一个bug (啊变量名写错了, 改一下改一下)
- 跑通了, 说一下思路吧 (简单解释了下代码思路, 理论应该是对的)
- 反问: 表现如何?
- 反问: 一共有几面
- 个人体会: 铁凉凉, 不用看了, 这也太难顶了, 手撕LSTM还问了一堆机器学习东西, 本来是准备的论文, BERT, word2vec, 结果都没问。这波死的及其惨烈~HMM, 新词发现, 这两不熟, 赶快补, 还有coding能力不够, 找机会多写代码!!!

二面

- 自我介绍, 聊项目论文。
- BERT和RNN之类的区别 (Auto Encoder与Auto regression, 吧啦吧啦)
- BERT的缺点, 什么情况无法处理, 双重否定, 数值可以处理吗 (应该不太明感)
- BERT的输入长度多少? (最大512) 超出怎么办 (截断, 数据预处理减少无关数据。实在太长直接上LSTM、Transforemer-XL)

- 实体抽取怎么做的，有了解吗？（没有具体做，但是大概知道上NER这一块的东西，其他不清楚了）
- 一个算法题，求一个大于等于输入的正整数的最小降序序列，比如，325->330, 4->4, 1234->2000, 4321->4321（20min没做出来，说了下思路，暴力递归）
- 个人体会：这个面试官又迟到10min，携程感觉不太守时，不喜欢，然后办公环境看起来也不太行，总的来说不太行。最后算法题没写出来，估计是凉了，不过这个面试官比上一个态度好很多。

腾讯

一面

腾讯会议面试（30min）

- 自我介绍，每个项目，一个个讲讲
- 论文1中样本怎么获取的（远程监督）继续深挖
- 论文2介绍一下（吧啦吧啦）
- BERT怎么做分类的说一下（feature based fine turning）
- 比赛做了什么，改动了什么，和第一差距在哪（吧啦吧啦）
- 多任务用和不用的区别在哪（pipeline存在误差传递，多任务2个loss加一起）
- 2个loss的超参数怎么获取（经验+网格搜索）这样可能很慢，贝叶斯优化了解吗（不是很了解，只记得是启发式算法）
- tf会用吗，和pytorch有什么区别（不会，一般都是pytorch，tf有session、graph什么的比较麻烦记得好像，其他不清楚了）
- 其他语言会吗（本科学过、捡起来很快）大数据开发会吗（不会，只听过）
- 树模型和深度学习的区别（只知道效果有差距，具体不太清楚）二者可以结合吗（不是很清楚）
- 树模型介绍一下（决策树、GBDT、XGBOOST、随机森林）核心思想介绍下（吧啦吧啦）
- Adaboost和GBDT区别（Adaboost对错误给更高的权重，GBDT多个子树加起来（不知道说的对不对））
- 树模型和机器学习比如LR有啥区别（树模型有比较好的可以解释性（这个答案好像不是很满意，应该是稀疏性的问题））

- 稀疏性问题怎么解决（变成dense向量，说了一些分布式表示和one-hot的区别（不知道答偏了没有））
- 推荐系统了解吗（不是特别的了解）想做哪一块（NLP这块吧）推荐转岗可以吗（可）
- 反问：业务有哪些呢
- 反问：技术栈要求
- 反问：面试不足的地方有哪些
- 个人体会：面试官语速快的飞起，有种压力面的感觉，前面说项目的时候打断了好几次，把我说话节奏都带的快了。还有就是coding能力不够，工业界的一些东西还是缺乏

二面

腾讯平台面试（1h20min）

- 自我介绍，问项目。
- attention知道吗，用过吗（BERT里用过，不是我写的）写一下公式（从attention说到self-attention）
- q、k、v是啥意思（解释了一下）qk可以用别的方法代替吗（可以，只要是计算距离的方法就行，比如说余弦距离）
- self-attention和attention的区别（说了一堆好像没答道点上，最核心应该是attention只计算q和v）
- BERT和其他序列模型区别（BERT不是序列模型，Auto-Encoder的模型，其他比如LSTM是Auto-Regression）
- 推荐算法了解吗（不太了解，只知道协同过滤）
- 机器学习了解吗（了解），LR损失函数写一下（吧啦吧啦写了下，最后y写错了，应该放log外面）
- FM了解吗？（不了解）Wide&Deep了解吗？（不了解，好像都是推荐系统的那边）
- XGBOOST说一下（说了下和GBDT区别，之类的）为什么二阶导数好（信息更丰富）二阶导数也会更好的拟合（对对对）
- LSTM解决了RNN什么问题（长依赖，梯度消失梯度爆炸之类的）
- LSTM还有一个什么特点？（啥。。。不知道了）做加法比做乘法快（喔喔喔，可能想问三个门并行化？没太理解）
- 转岗推荐可以吗（可）

- 说说对推荐系统理解（只知道协同过滤，时间一致性巴拉巴拉）
- Embedding的理解，说说（表示学习巴拉巴拉说一堆）
- 写一个二叉树的后序遍历，非递归，手写（双栈法，输出的时候有点小问题）
- 算法题手写：最大子段和（暴力法）
- 反问：面试表现
- 反问：推荐的技术栈
- 个人体会：体验很好，就是写代码有点紧张。腾讯的速度还是很快的，不得不感叹。

HR面

腾讯会议面试（18min）

- 自我介绍，实习时间，信息确认
- 挑一个项目说一下
- 项目中的收获是什么
- 说2个项目中的难点，怎么解决的
- 希望实习获得什么
- 别的公司进度
- 有重大疾病吗
- 如何和别的部门协调资源
- 实习和实验室做项目区别
- 给自己之前2个面试打分
- 反问：表现怎么样
- 反问：base地（北京、深圳）
- 个人体会：腾讯速度真的快的飞起，结束了就云证。昨天2面今天就HR，不得不说太快了。HR态度很客气，体验很好