

## 实习经历重点描述

### 自我介绍部分

面试官，下午好，我是来自大连理工大学的硕士研究生，今年研三，我叫纪杰周，很高兴能参加咱们公司的面试，我之前有过三段实习，第一段实习是在千里马网信科技有限公司，这家公司主要负责招投标信息处理的，然后我在这家公司主要工作就是基于自然语言处理来处理一些招投标文本，第二家实习公司是 MetaApp，这家公司是一家游戏公司，在那边主要是负责构建基于 TypeScript 语言的代码生成模型与代码补全模型，最后一段实习实在美团，主要做的就是数据工程实验与业务场景下 LLMs 的推理能力建设及应用。

### 美团

◎ **问题背景：**业务方现有的表格查询方式主要依靠点击式操作，对于新员工来说学习成本较高，需要时间熟悉各种查询条件和操作步骤，基于此，业务方那边希望我们这边能够提供一个能处理 Text2SQL 的大模型，方便新员工通过自然语言进行表格查询。

**实习描述：**针对招聘场景构建 Text2SQL 任务，首先对业务方提供的 SQL 代码与查询需求进行了简单的分类，（查询需求共分成三种，单表单值，单表多值，多表多值），基于此实现了查询需求的自动化生成并结合 GPT4 构造微调指令数据集（3000 条），并随机抽样检查了数据质量，随后跟之前微调的 20w 融合做微调，微调后模型在自建的 benchmark 上的 pass@1 达到了 0.8，随后，让业务方那边去评估生成质量，后用拒绝采样的方法，又添加了 2000 条数据集，最终在自建 benchmark 上的 pass@1 达到了 0.95。

◎ **问题背景：**去年以来，很多论文提出，使用少量高质量的微调数据可以取得良好的效果。此外，一些研究者也公开了一些高质量数据筛选方法。基于此，我们希望在现有的数据质量过滤方法基础上，进一步优化数据筛选流程。因为这样做，一方面可以节省训练资源，另一方面也可以探究下模型的上限。

**实习描述：**我们这里有 2kw 的指令数据集，其中包括业务和公开数据，首先说一下我们原先数据质量筛选 pipeline，先利用 k-means 对数据进行分簇，分了 1000 簇，然后再结合我们的数据质量评分模型，对每簇里的数据进行评分，取质量评分靠前的指令数据，新的数据筛选 pipeline 就是在这个基础上，又做了两层筛选，第一层是计算完数据质量分数之后，然后计算每条数据的平均交叉熵损失，去掉较大的或者较小的，然后在计算每簇数据的 IFD 值，最后采样出数据集。

- **K-means** 是无监督学习算法，不需要事先标记的数据。它通过迭代的方式将数据聚类成 K 个簇，使得每个簇内的点尽可能相似。
- **KNN** 是有监督学习算法，需要事先标记的数据。它通过计算待分类点与训练数据集中已知类别的点的距离，将其归为距离最近的 K 个点的类别。

◎ **问题背景：**建设模型的推理能力，一方是原模型的推理能力在业务场景重表现较差（具体表现为无 COT 对测评结果影响不大，而且当对于触发 COT 也不稳定，就算触发成功，但是其逻辑较为混乱），另一方面是，有部分业务场景的任务需要中间的推理过程来辅助做处理下游任务。（例如：商品推荐，退款解释，等等）

**实习描述：**首先我们对业务场中的任务进行了分类，即：“重结果”与“重过程”，这里有必要解释下两者，重过程就是，这个任务，强调的是通过一步一步的逻辑推理，并最终得出答案，强调的中间的推理过程的正确性，类似于数学题，也可以想这也就是思维链，像我们业务场景中的要点打判、上报异常等任务。重结果就是，模型不仅仅给出了问题的结果，更加强调的是，为什么会给出这个结果，即模型对它给出这个结果的一个合理解释，这就是重结果，像我们的商品推荐、对话总结。在此基础上，分别结合 orca1 和 orca2 的思想去构造业务场景下的增强数据，简单说一下 orca1 与 orca2 吧

Orca1 指出，拿推理数据区微调小模型时，小模型往往只能模仿推理的输出风格，而非推理过程，所以导致它们面对复杂或模糊的咨询经常会犯错或给不出答案。Orca1 核心是解释调优。让 GPT-4 通过“顺

序思考”或”向小学生那样进行解释”之类的提示来揭示其思维过程，这样 GPT-4 的推理过程就变得更加透明。这样丰富和结构良好的反应允许调整小模型来模仿 GPT-4 的思维过程。小模型去学习 GPT-4 的解释痕迹、逐步的思维过程、复杂的指令等来完成各种任务。然后 orca2 就是在 orca1 的基础上精心设计推理策略以适应特定任务，即用什么策略解决什么问题。（3w 业务增强数据）

后对 Ziya-Llama-13B 进行解释调优，同时设置 Magic-Prompt 激发模型针对特定问题的推理能力。通过分析错误 Case 的推理过程不断修正增强数据集，最终使得微调后的 Ziya-Llama-13B 在业务榜单提升约 10 个 pp，且推理能力提升明显。

◎ **问题背景：**现有的模型推理能力存在一些缺陷，例如，偶尔会出现推理过程与答案不一致的情况，或者答案正确但推理过程很差的情况。仅仅用答案的正确性来衡量模型推理过程的合理性存在局限性。为了解决这个问题，我们提出了建设 Critical-model 的方法，以更合理地评估模型推理过程的合理性。

**实习描述：**在建设 Critical-model，我们做了大量的调研，针对我们的业务场景，要思考的问题就是，对文本的评估，应该适合逐点扣分还是打总分？后来我们在重过程的任务上找扣分点，对于重结果的任务进行打总分，即：

重过程：扣分点，错误定位，建议的修改意见

重结果；打总分，不足之处，建议的修改意见

确定好任务之后，我们就去收集了类似数据，一方面有开源的，另一方面我们基于业务也构造了一部分，首先我们挑选了一个业务数据子集，在这个子集的基础上去询问各类开源的模型，然后用 gpt-4 或者外包去针对模型的回答进行打分并给出修改意见，最后共得到 2w 的数据。训练完成之后，又跟其他的模型做了比较，发现我们的模型更好。（与人工打分更加正相关）

## MetaApp

◎ **实习描述：**代码数据的预处理流程如下：（首先要保证所收集到的代码数据都提前做好了分块，预处理阶段只需要做异常检测）：去重（LSH, MIN-hash）、多余格式去除，规范代码、对缺少注释的代码块的添加注释、代码异常检测，去除运行异常的代码

◎ **问题背景：**提供代码生成模型，供游戏开发部门使用。

**实习描述：**训练策略共分成 继续预训练+SFT、继续预训练+SFT(Lora)、直接 SFT、直接 SFT(Lora)，继续预训练数据共 15GB（其中 TS:其他语言=1:1），2w 条指令数据集（TS【业务+通用】+其他），4000 条训练 RM 模型的数据。

阶段	MXEval-pass@1	业务榜单Benchmark-pass@1
gpt3.5	0.6	0.7
starcoderbase	0.45	0.5
starcoderbase+pretrain	x	x
starcoderbase+SFT	0.55	0.6
训练RM	0.58 (如果以RM最高分参考来进行测评)	0.63
PPO	0.66	0.84

◎ **问题背景：**在测评中发现了灾难性遗忘这个问题，然后后面对其进行修复。

**实习描述：**重点参考了 InstrucGPT 的第三阶段 ppo 损失函数设置，强行保持微调模型离参考模型不那么遥远，最后发现有用。