

Criteria Learning Proposal及论文大纲

Meta Prompting via Criteria Generation

Motivation: 之前的工作发现criteria不仅有助于提升反思的质量，还有助于提升模型的生成回复的质量。其有潜力成为一种meta-prompting的策略。

1. 首先验证criteria可以提升回复的质量
2. 在这篇工作中，我们首先希望探讨一个高质量的Criteria应该具有什么样的特性：
 1. criteria的数量
 2. criteria的组织结构：层次化的Criteria - 类似HD-Eval的工作
 3. criteria的格式：有score rubric, 无score rubric, mix
 4. 甚至可以对比一些已有工作的criteria构建方法，如prometheus和HD-Eval
3. 在以上探讨的基础上，我们分析使用criteria提升数据合成质量的可能性，进而有助于构造大规模高质量合成SFT数据
4. 在上述基础上，我们prompt GPT-4构建了一个高质量的SFT数据集
5. for RAG ...?
6. for Multi-agent ...?

我们的方法和已有的提升数据质量的方法，如Self-Critique是正交的，可以作为一种通用的prompting技巧——meta prompting

1. Criteria提升数据的质量

首先验证默认的criteria是可以提升数据质量的。这里我们采用类似Auto-J的评估维度（无score rubric, 无层次化，默认5条criteria）来评估。具体的实验设置：

"C:\Users\18811\OneDrive\Doc\Document\my typora notes\critiera-primlinary-study-proposal.pdf"

2. 分析最优Criteria的配置

从三个角度分析一个好的criteria应该是什么样的：

1. criteria的数量
2. criteria的组织结果
3. criteria的格式

3. baseline和已有的提升数据质量方法的对比

1. self-critique @高学弟
 1. 采用多轮自我反思流程，比如从1-5轮，每一轮都要模型先生成critique，在依据critique生成 revised response
2. rejection sampling: <https://mingchao.wang/glk6wBrl/> @明学弟
 1. 需要留意采样阶段的temperature系数，可以参考上面链接

4. 构建高质量Instruction-Following数据集

1. 在上述实验结论的基础上，我们构建一个现有的高质量Instruction Following SFT数据集
2. 微调模型并分析模型的性能
3. 分析训练阶段联合训练criteria对于生成高质量回复的影响

5. TODO List

5.1 追踪和调研当前的相关论文

1. Prometheus系列论文
 1. <https://arxiv.org/abs/2310.08491>
 2. <https://arxiv.org/abs/2405.01535>
2. Auto-J: <https://arxiv.org/abs/2310.05470>
3. UltraFeedback: <https://arxiv.org/abs/2310.01377>
4. Self-Critique相关论文：
5. 提升回复质量的相关论文：
6. HD-eval: <https://arxiv.org/pdf/2402.15754>

5.2 准备github repo

- 维护prompt和baseline数据构造流程