Paper Reading 7

沙之洲 20200124080

Constitutional AI：

1 这篇工作在构建训练数据集的很重要的一步是，首先让 LLM 生成对危险问题的回答，然后根据人工给定的 principles 对给定的回答进行评判。但是有一种观点认为，LLM 生成回答的一个基本原则是为了保证生成的语句通顺连贯，我们该如何区分模型是在做有逻辑生成，还是仅仅是依靠“语感”在做连贯性生成？

2 这篇工作提出了一种基于 AI 的 self-supervised 的来 follow 认为给定的 helpfulness 和 harmlessness principles。是否考虑可以未来将 principles 的种类进行扩展，根据客户的需求 finetune 出具有特定风格的模型？比如说，小说创作家需要模型给出激进新颖的想法，而企业管理者往往更喜欢稳健实用的建议，可以根据这两种不同的需求设计对应的 principles，让模型能够更好满足他们的需要。

AnnoLLM：

1 这篇工作的核心思想是，针对一个标注问题，首先让 LLM 给出 explanation，然后基于 explanation 给出 Chain of thought 的 prompt，再让 LLM 生成最终的标注结果。也就是说，同一个LLM 在标注生成的过程中在扮演着不同的角色，是否可以让多个 LLM 副本同时生成标注结果，以投票的方式决定最终的标注结果？也即，让 LLM 像团队一样进行合作，给出比单个 LLM 自己标注结果更好的结果？

2 将上述 LLM 之间合作的思想进一步延伸。是否可以让每一个 LLM 副本尽可能去拟合一个领域的知识，让后通过 actor-critic 的方法，让擅长每个子领域的 LLM 作为 critic，通过问答学习的方式将所有子领域的知识整合到一个更大更统一的 LLM 上？