

小样本提示学习篇

来自：AiGC面试宝典

宁静致远

2024年01月27日 19:14



- 小样本提示学习篇
 - 一、什么是Zero-shot提示方法？
 - 二、什么是Few-shot提示方法？
 - 三、阐述One-shot和Few-shot提示策略及其应用场景？
 - 四、什么是逐步Zero-shot
 - 五、定义Zero-shot-CoT提示策略并描述其应用方法？
 - 六、解释Few-shot-CoT提示策略及其实际使用方式？
 - 七、Few-shot-LtM策略包含哪些主要阶段及其职责？
 - 致谢

一、什么是 Zero-shot 提示方法？

Zero-shot 提示方法是指给模型提供一个不属于训练数据的提示，但模型可以生成你期望的结果的方法。这种方法使得大型语言模型可以用于许多任务。例如，你可以给模型一个提示，让它翻译一句话，或者给出一个词的定义，或者生成一首诗。最简单的提示工程方法就是通过输入一些类似问题和问题答案，让模型参考学习，并在同一个 prompt 的末尾提出新的问题，zero-shot 可以理解为：不给大模型任何的提示，直接提问，让大模型自己做决策。

二、什么是 Few-shot 提示方法？

Few-shot 提示方法是指给模型提供一些示例或上下文，来引导模型更好地完成任务的方法。这些示例或上下文可以作为模型的条件，来影响后续的输出。例如，你可以给模型一些情感分析的示例，让它根据文本判断情感倾向，或者给模型一些编程任务的示例，让它生成代码片段。可以简单理解为：在提问之前，先给大模型一个示例和解释让它学习和模仿，从而在一定程度上赋予它泛化能力。

三、阐述 One-shot 和 Few-shot 提示策略及其应用场景？

One-shot 或者 Few-shot 提示方法的思想：最简单的提示工程的方法就是通过输入一些类似问题和问题答案，让模型参考学习，并在同一个 prompt 的末尾提出新的问题，以此提升模型的推理能力。One-shot 和 Few-shot 提示策略及其应用场景如下：One-shot 提示策略是指只给模型一个示例或上下文的方法。这种方法适用于一些相对简单或常见的任务，或者一些模型已经有了很强的涌现能力的任务。例如，你可以给模型一个新词和它的定义，然后让它用这个新词造句。Few-shot 提示策略是指给模型多个示例或上下文的方法。这种方法适用于一些相对复杂或特殊的任务，或者一些模型需要更多的引导和调整的任务。例如，你可以给模型几个不同类型的笑话，然后让它根据一个关键词生成一个新的笑话。具体的应用来说，Few-shot 提示方法并不复杂，只需要将一些类似的问题+答案作为 prompt 的一部分进行输入即可。当需要输入多段问答作为提示词时，以 Q 作为问题的开头、A 作为回答的开头（也可以换成“问题”、“答案”），并且不同的问答对话需要换行以便于更加清晰的展示，具体方法是通过转义符+换行来完成。

四、什么是逐步 Zero-shot

逐步 Zero-shot 是指一种利用大型语言模型进行推理的方法，它通过让模型自动生成多个中间步骤或过程，然后再根据这些步骤或过程来生成结果。逐步 Zero-shot 可以让模型更好地理解问题的含义和范围，更接近人类的思考方式。在论文 *Large Language Models are Zero-Shot Reasoners* 中，提出了“逐步 Zero-shot”这一思想。。在第一个阶段，模型首先对问题进行拆分并分段解答问题（称为 Reasoning Extraction）。然后，在第二阶段，模型再进行答案的汇总（称为 Answer Extraction）。

五、定义 Zero-shot-CoT 提示策略并描述其应用方法？

Zero-shot-CoT 提示策略是指一种利用 Zero-shot 思维链进行推理的方法，该方法借助思维链（也被称为思考链，Chain of Thought, CoT）提示法来解决问题。它通过在问题的结尾附加“Let's think step by step”这几个词，来引导模型生成一个回答问题的思维链。Zero-shot-CoT 提示策略可以让模型更有效地利用其通用知识和涌现能力，同时避免过拟合或灾难性遗忘。

Zero-shot-CoT 提示策略的应用方法如下：

1. 首先，明确你想要解决的问题或任务，以及你期望的输出或回答的格式和内容。
2. 然后，将你的问题或任务用自然语言描述出来，并在结尾加上“Let's think step by step”这几个词，作为输入或查询给模型。
3. 最后，观察模型生成的输出或回答，看是否符合你的期望和需求。如果不满意，可以尝试修改你的输入或查询，或者使用其他控制参数来影响模型的行为。

例如，你可以使用 Zero-shot-CoT 提示策略来解决一个简单的数学应用题：

提示 我去市场买了 10 个苹果。我给邻居 2 个苹果，给修理工 2 个苹果。然后我又买了 5 个苹果，并吃了 1 个。我还剩下多少个苹果？ Let's think step by step.

输出 首先，你开始有 10 个苹果。你送走了 2 个苹果给邻居和 2 个苹果给修理工，所以你还剩下 6 个苹果。然后你又买了 5 个苹果，所以现在你有 11 个苹果。最后，你吃了 1 个苹果，所以你还剩下 10 个苹果。

六、解释 Few-shot-CoT 提示策略及其实际使用方式？

Few-shot-CoT 提示策略是指一种利用 Few-shot 思维链进行推理的方法，它通过给模型提供一些手动设计的中间步骤或过程，来影响模型的输出。Few-shot-CoT 提示策略可以让模型更专注于新任务或领域的特点，提高模型的性能和准确度。其核心思想是通过编写思维链样本作为提示词，让模型学会思维链的推导方式，从而更好的完成推导任务。需要在提示样本中不仅给出问题的答案，还同时需要给出问题推导的过程（即思维链），从而让模型学到思维链的推导过程，并将其应用到新的问题中。Few-shot-CoT 的方式虽然有效，但是并不是很稳定，如果想要得到稳定的正确答案，可能需要更高阶的提示方法。

Few-shot-CoT 提示策略的应用方法如下：

1. 首先，明确你想要解决的问题或任务，以及你期望的输出或回答的格式和内容。
2. 然后，将你的问题或任务用自然语言描述出来，并在前面加上一些与任务相关的中间步骤或过程，作为输入或查询给模型。
3. 最后，观察模型生成的输出或回答，看是否符合你的期望和需求。如果不满意，可以尝试修改你的输入或查询，或者使用其他控制参数来影响模型的行为。

七、Few-shot-LtM 策略包含哪些主要阶段及其职责？

Few-shot-LtM 策略是一种利用最少到最多的提示序列，来让大语言模型逐步增加推理难度和深度的方法。它可以让模型更有效地利用其通用知识和涌现能力，同时避免过拟合或灾难性遗忘。Few-shot-LtM 策略包含以下两个主要阶段及其职责：

1. 第一阶段是将问题分解为子问题，即根据问题的复杂性和模型的能力，将问题拆分成若干个更简单或更具体的子问题。这些子问题可以建立在彼此之上，也可以相互独立。这个阶段的职责是为模型提供一个清晰和合理的思路，让模型能够逐步接近最终的目标。
2. 第二阶段是逐个解决子问题，即根据第一阶段得到的子问题序列，依次给模型提供相应的提示，让模型生成每个子问题的答案。这些答案可以作为下一个子问题的输入或条件，也可以直接作为最终答案。这个阶段的职责是为模型提供一个有效和灵活的引导，让模型能够产生合理和正确的输出。