

小组报告

LLMs 大语言模型

GPT 发展历程

GPT（Generative Pre-trained Transformer）是自然语言处理领域中的一种预训练语言模型。下面是 GPT 发展历程的概述：

● 2017 年：Transformer 模型的提出

2017 年，谷歌的研究人员提出了 Transformer 模型，这是一种基于自注意力机制的神经网络结构，用于将序列数据转换为另一个序列数据。Transformer 模型解决了长文本序列处理的问题，成为了自然语言处理领域的重要突破。

● 2018 年：GPT-1 的推出

2018 年，OpenAI 发布了第一个基于 Transformer 模型的预训练语言模型，称为 GPT-1。GPT-1 遵循“预训练-微调”范式：对于大规模无标注语料库的学习，使用单向生成式目标函数进行预训练；随后针对不同自然语言处理任务，如机器翻译、文本摘要等，使用有标注数据进行有监督的微调使得模型能更好地适应下游任务。

● 2019 年：GPT-2 的发布

2019 年，OpenAI 发布了更加强大的预训练语言模型 GPT-2。GPT-2 更加侧重于 zero-shot 设定下语言模型的能力，即模型在下游任务中不进行微调和参数优化，而是根据给定的指令自行理解并完成任务。GPT-2 并没有对 GPT-1 的模型架构进行创新，而是在 GPT-1 的基础上引入任务相关信息作为输出预测的条件；并采用了更大的数据集。

● 2020 年：GPT-3 的发布

2020 年，OpenAI 发布了 GPT-3，使用了与 GPT-2 相同的模型和架构。为了探索模型规模对于性能的影响，一共训练了 8 个不同大小的模型，并将最大的具有 1750 亿参数的模型称为 GPT-3。模型本身规模大，参数量众多，具有 96 层 Transformer Decoder Layer，每一层有 96 个 128 维的注意力头，单词嵌入的维度也达到了 12,288；另一方面是训练过程中使用到的数据集规模大，达到了 45TB。

从 GPT-1 到 GPT-3，模型训练的规模越来越大，参数越来越多，模型在自然语言任务上的表现也更佳。GPT-3 在问答、文本摘要、机器翻译和语言生成等多个任务上均展现出非常优异的性能，在 Few-shot，One-shot 和 Zero-shot 等设置下的任务表现也得到了显著的提升。

模型	发布时间	层数	头数	词向量长度	参数量	预训练数据量

GPT-1	2018 年 6 月	12	12	768	1.17 亿	约 5GB
GPT-2	2019 年 2 月	48	-	1600	15 亿	40GB
GPT-3	2020 年 5 月	96	96	12888	1,750 亿	45TB

● 2022 年 12 月 : ChatGPT 推出

2022 年 11 月 30 日，OpenAI 推出全新的对话式通用人工智能工具 ChatGPT。ChatGPT 表现出了非常惊艳的语言理解、知识推理、生成能力，它可以很好地理解用户意图，做到有效的多轮沟通，并且回答内容完整、有逻辑、有条理。ChatGPT 上线后，5 天活跃用户数高达 100 万，2 个月活跃用户数已达 1 个亿，成为历史上增长最快的消费者应用程序。

● 2023 年 3 月 : GPT4 推出

GPT4 是一个多模态模型，接受图片和文字的输入，并且输出文字。它可以理解图片内容，并且在文字输出结果上准确度和性能更好；可定制 AI 的风格和行为，而不是像 ChatGPT 那样具有固定冗长的语气和风格；在一些专业的学术考试中达到人类前 10% 的水准，远超 ChatGPT。[1]

ChatGPT 的核心技术

● 实现基础 : Transformer 架构[2]

原始的 Transformer 模型包含两个关键组件：编码器和解码器。编码器用于将输入序列映射到一组中间表示，解码器则将中间表示转换为目标序列。编码器和解码器都由多层的注意力模块和前馈神经网络模块组成。其中自注意力模块可以学习序列中不同位置之间的依赖关系，即在处理每个位置的信息时，模型会考虑序列中其他所有位置上的信息，这种机制使得 Transformer 模型能够有效地处理长距离依赖关系。

在原始 Transformer 模型基础上，相继衍生出了三类预训练语言模型：编码预训练语言模型、解码预训练语言模型和编解码预训练语言模型，GPT 系列属于解码器的预训练模型。

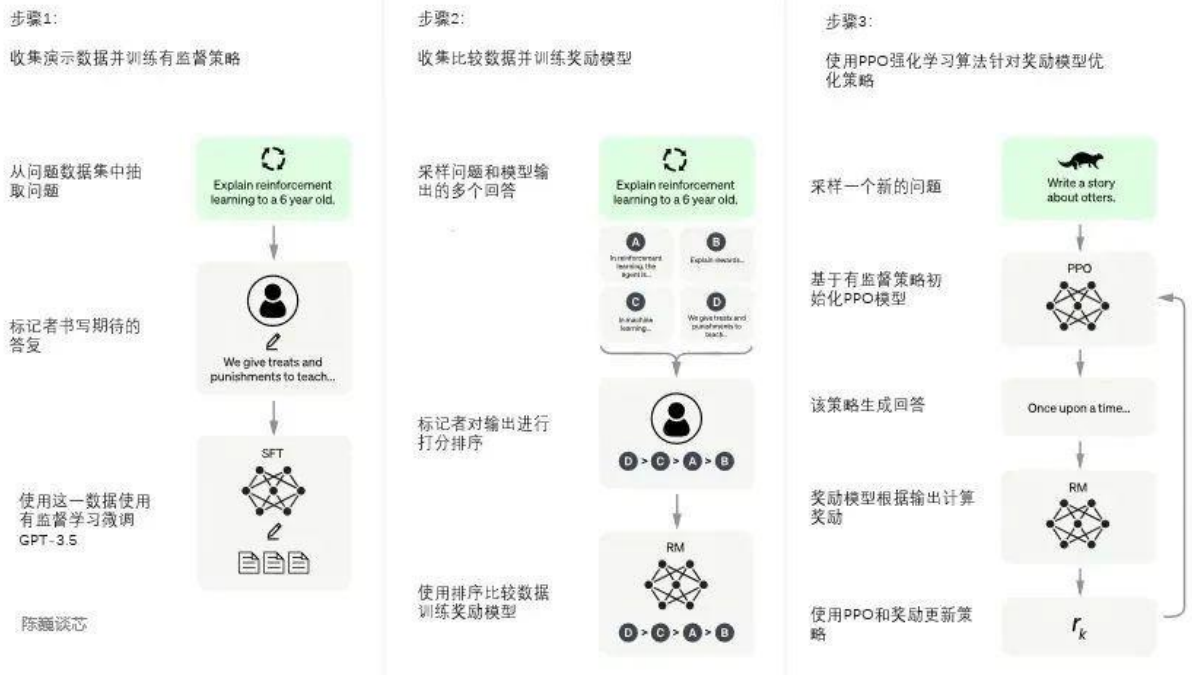
● 提示学习和指令微调

ChatGPT 所用到的指令学习数据集的构造方法和训练方法与 InstructGPT 大致相同 [3]，因此我们介绍 InstructGPT 构造“指令”数据集的细节。

InstructGPT 的“指令”数据集由两部分构成，其中一部分收集于全球用户使用 OpenAI

的 API 后的真实人机交互数据（这些数据在使用之前都经过了信息去重和敏感信息过滤）；另一部分数据则来自于人工标注。这些人工标注的数据总共分为三类，其一是为了增加数据集中任务的多样性，由标注人员写出任意任务的“指令”；其二是小样本（Few-shot）数据，即标注人员写出的“指令”和一些对应的问答对，用于训练模型的小样本学习（Few-shot learning）能力；其三是模仿 OpenAI API 已有的用例写出的相似“指令”。这些数据包含了语言模型中常见的任务类型（生成、问答、聊天、改写、总结、分类等），而生成任务在其中的占比最大。

InstructGPT 通过在构造的”指令”数据集上进行有监督微调（Supervised fine-tuning, SFT）和基于人工反馈的强化学习（Reinforcement Learning from Human Feedback, RLHF）以使模型与人类需求对齐。



思维链 Chain of Thought

[4]通过在小样本提示学习的示例中插入一系列中间推理步骤，有效提升了大规模语言模型的推理能力。针对零样本场景[5]，利用推荐关键词“Let’s think step by step”(让我们一步一步思考)生成中间步骤的内容，从而避免了人工撰写中间步骤的过程。

(a) Few-shot

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The answer is 8. ✗

(b) Few-shot-CoT

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are $16 / 2 = 8$ golf balls. Half of the golf balls are blue. So there are $8 / 2 = 4$ blue golf balls. The answer is 4. ✓

(c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(Output) 8 ✗

(d) Zero-shot-CoT (Ours)

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓

基于人类反馈的强化学习

RLHF 是 ChatGPT/InstructGPT 实现与人类意图对齐，即按照人类指令尽可能生成无负面影响结果的重要技术[6]。该算法在强化学习框架下实现，大体可分为以下两个阶段：

- 奖励模型训练

该阶段通过对于那些涉及偏见的生成内容人工赋低分的方法，鼓励模型不去生成人类不喜欢的内容，尽可能获取拟合人类偏好的奖励模型。具体做法是首先生成一批候选文本，然后人类给出倾向排序，模型进而拟合倾向性标签，完成训练。

- 生成策略优化

给定习得的奖励模型，ChatGPT/InstructGPT 的参数将被视为一种策略，在强化学习的框架下进行训练。当前策略根据输入做出回应，奖励模型针对回复计算奖励，反馈回当前策略用以更新。此外，为防止上述过程的过度优化，损失函数同时引入了词级别的 KL 惩罚项。此外，为了避免在公开 NLP 数据集上的性能退化，策略更新过程兼顾了预训练损失。

ChatGPT 带来的风险和挑战

- 滥用风险

用户可能将 ChatGPT 应用于不当领域，例如学生利用其完成作业、研究人员利用其代笔学术写作、不法分子利用其制造假新闻等。针对这一问题，有人提出了对 AI 生成文本的鉴别技术，以降低这一类不当应用的潜在风险。

- 错误信息风险

在有监督的语言模型任务中，人类只是起到了纠正作用。受限于纠正数据的有限，或是有监督任务的误导，导致 ChatGPT 给出一些荒谬的输出。ChatGPT 的流行会在某种程度上增加用户对它的信任，可能使用户被更多未经验证的信息误导。如果将生成的信息应用在比较重要的场合，例如急救和法律援助，可能会出现未知的风险。

● 心理健康风险

大多数人没有对人工智能的基本认识以及足够的情感和伦理素养，因此可能会产生情感问题和心理健康风险。一方面，不加限制的 ChatGPT 可能输出的暴力色情等信息会影响使用者的心理健康；另一方面，一些用户潜意识中会将交互式对话 AI 当做人类来交流，对其可能会产生依赖性 or 成瘾性。

● 垄断风险

ChatGPT 对训练数据、算力和人力的要求都很高，需要大量经费投入，如此高开发门槛的技术可能被财力雄厚的大公司垄断，进而影响该领域相关的企业间的公平竞争。另外，由于很多应用已经建立在 ChatGPT 之上，一旦形成垄断，这些应用就都必须依附于垄断者，其商业或研究活动就会受到制约。

展望

BERT-Large(2018)有 3.55 亿个参数，GPT-2(2019 年初)达到 15 个 b, T5(2019 年底)进一步延伸到 11B, GPT-3(2020 年年中)最终达到 175B，而 GPT3.5，GPT4 的参数数量只会更加恐怖。语言模型大小的进步明显超过了 GPU 内存的增长。这意味着对于 NLP 来说，“令人尴尬的并行”的时代即将结束，模型并行化将成为研究 SOTA 语言模型不可或缺的一部分。

将人类大脑与我们的语言模型进行比较[7]:一个典型的人类大脑有超过 100 万亿个突触，这比 GPT-3 175B 模型大了三个数量级。OpenAI 仅用了一年零一个季度的时间就将 GPT 模型的容量从 15 亿增加到 1750 亿，提高了两个数量级，因此更大的语言模型很有希望。

参考资料：

主要参考：车万翔等.ChatGPT 调研报告[R].哈尔滨工业大学自然语言处理研究所，2023.

[1]<https://openai.com/research/gpt-4>

[2]<https://zhuanlan.zhihu.com/p/338817680>

[3]<https://openai.com/blog/chatgpt>

[4] WEI J, WANG X, SCHUURMANS D, et al. Chain of thought prompting elicits reasoning in large language models[J]. ArXiv preprint, 2022, abs/2201.11903(引用页: 32, 33)

[5] KOJIMA T, GU S S, REID M, et al. Large language models are zero-shot reasoners[J]. ArXiv preprint, 2022, abs/2205.11916 (引用页: 33)

[6] OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback[J]. ArXiv preprint, 2022, abs/2203.02155 (引用页: 11, 33, 48).

[7]<https://news.ycombinator.com/item?id=23345379>