



AI工具辅助学术写作 ——以ChatGPT为例

Academic Writing with ChatGPT

Presenter: 聂耳聪 Ercong Nie

2023.09.25
湖北大学

目录

- 一、什么是ChatGPT？
- 二、ChatGPT如何辅助学术写作？
- 三、案例分析
- 四、其他学术写作润色辅助工具
- 五、交流讨论

介绍

聂耳聪

- 德国**慕尼黑大学**信息与语言处理中心博士生，同时隶属于慕尼黑机器学习研究中心
- 研究领域为计算语言学、自然语言处理
- 本科(**德语**+金融, 上海交通大学)、研究生(**计算语言学**+计算机科学, 德国慕尼黑大学)
-

目录

一、什么是ChatGPT？

二、ChatGPT如何辅助学术写作？

三、案例分析

四、其他学术写作润色辅助工具

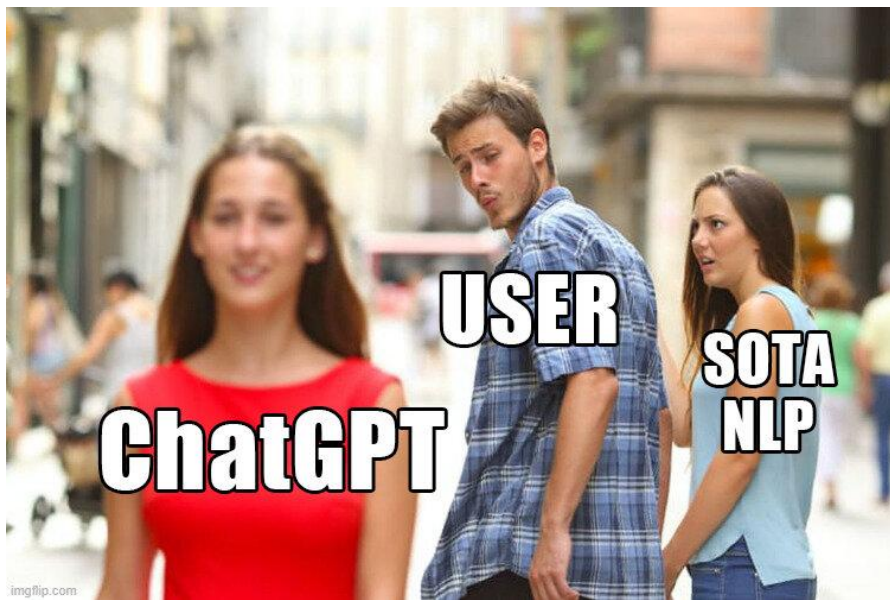
五、交流讨论

什么是ChatGPT？

- ChatGPT, 全称“**聊天生成预训练转换器**” (英语: Chat Generative Pre-trained Transformer)
- OpenAI开发的人工智能技术驱动的自然语言处理工具, 于2022年11月推出
- 能够基于在预训练阶段所见的模式和统计规律, 来生成回答, 还能根据聊天的上下文进行互动, 真正像人类一样来聊天交流, 甚至能完成撰写邮件、视频脚本、文案、翻译、代码, 写论文等**复杂自然语言处理任务**。
- 正在席卷各个垂直领域, 标志着人工智能**大模型** (LLM, Large Language Model) 时代的来临
- 人工智能工程项目——应用领先于理论: 仍有大量需要学界业界进一步研究的问题, 包括语言能力、推理能力、可解释性、领域大模型等。

什么是ChatGPT？

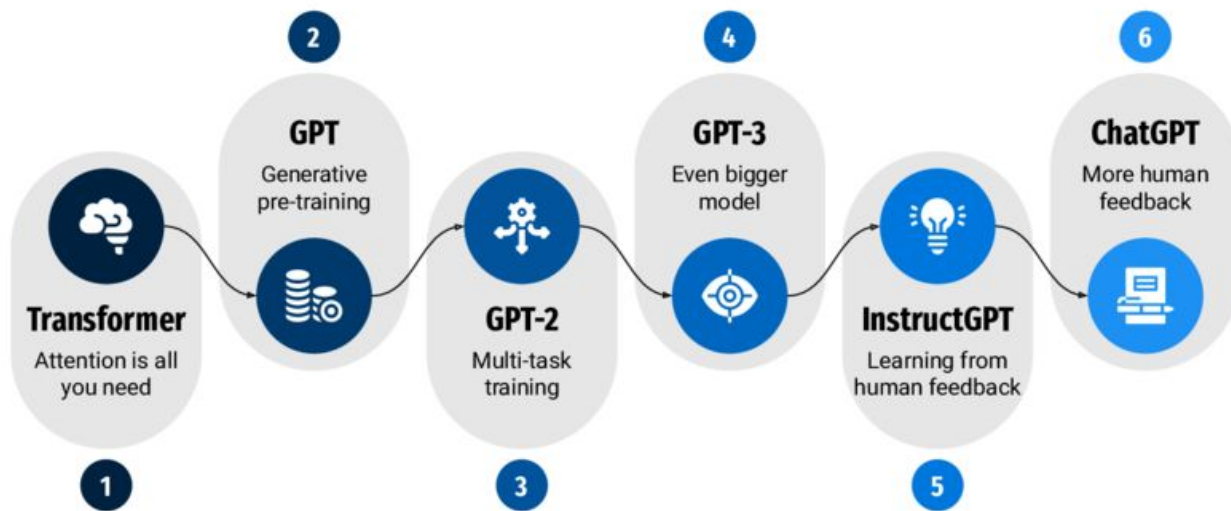
对于自然语言处理领域的用户来说，ChatGPT的出现就好像是：



但是究竟ChatGPT是什么？
它为什么行？

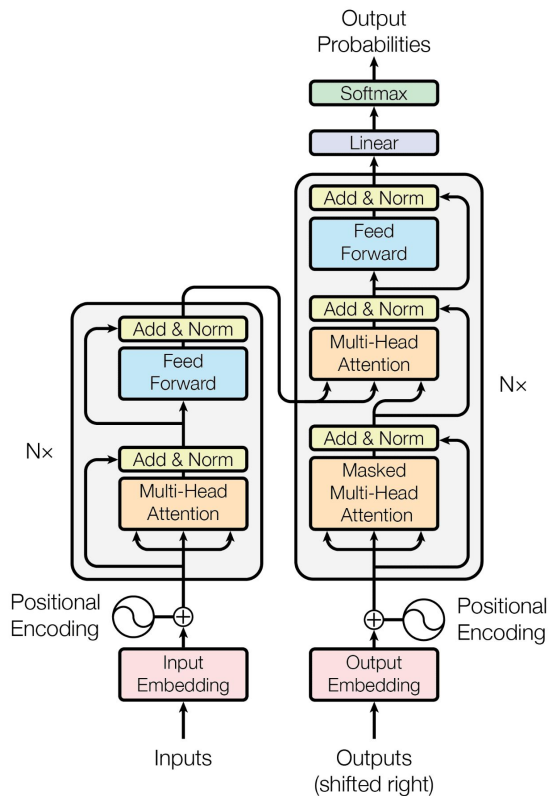
ChatGPT的发展

Evolution from Transformer architecture to ChatGPT



Source: Kocoń, Jan, et al. "Chatgpt: Jack of all trades, master of none." arXiv preprint arXiv:2302.10724 (2023).

ChatGPT的核心(1): 转换器(Transformer)结构



- 由编码器和**解码器**两部分组成
- ChatGPT**预训练阶段**的模型结构

Source: Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

ChatGPT的核心(2): 指令微调——加入人工反馈的强化学习

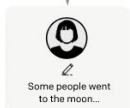
Step 1

Collect demonstration data, and train a supervised policy.

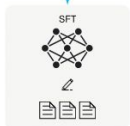
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

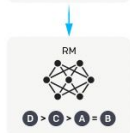
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

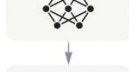
A new prompt is sampled from the dataset.



The policy generates an output.



The reward model calculates a reward for the output.

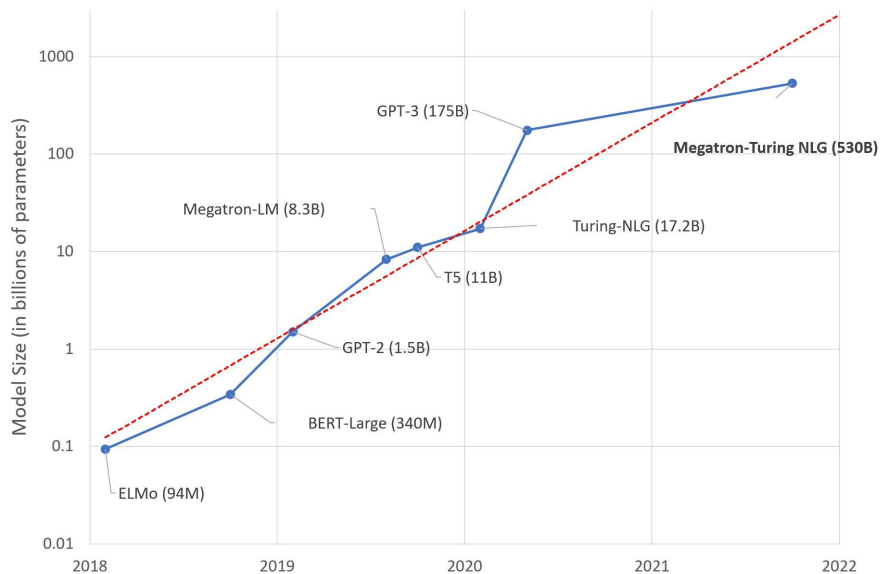


The reward is used to update the policy using PPO.

- ChatGPT对话能力的来源
- 需要海量的人工标记数据

Source: Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

大模型时代

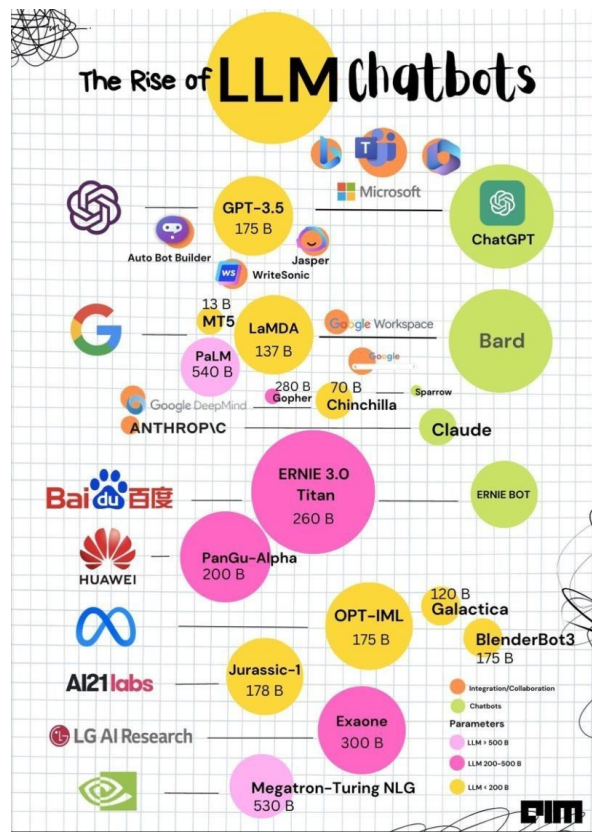


Sources:

<https://huggingface.co/blog/large-language-models>

<https://www.linkedin.com/pulse/talk-nerdy-me-how-large-language-models-changing-game-evan-kirstel/>

Sept. 2023



目录

一、什么是ChatGPT？

二、ChatGPT如何辅助学术写作？

三、案例分析

四、其他学术写作润色辅助工具

五、交流讨论

ChatGPT辅助学术写作的基本原理

- 了解大模型特点和本质:按照与AI Agent交互的方式使用大模型,“驯服”ChatGPT,从而满足用户的个性化需求,包括学术写作辅助
- ChatGPT辅助学术写作的能力来源:
 - **“学术能力”**:预训练语料中学术论文语料库占有相当大的比例(比如arxiv语料库),术语表达、学术写作风格能力的获取。
 - **“写作能力”**:指令微调阶段有多种写作能力的锻炼,比如文本总结、生成、续写、理解等。
 - **“知识能力”**:预训练语料中的**大规模知识库**,比如维基百科等
- 大模型使用核心:**提示**(Prompt)学习

ChatGPT辅助学术写作的策略

- 不能指望大模型进行“**一劳永逸**”的学术写作辅助
- 是自己的论文, 而不是大模型的论文: idea、思路、框架是自己的
- 战略上自我原创, 战术上AI辅助
 - **局部润色**
 - 文本**摘要生成**(辅助写一些总结性的话语、摘要、题目等)
 - 研究背景、比较general的话语辅助生成
 - 表格分析(GPT-3.5达不到学术写作要求, GPT-4可以)
- 专业性特别强的部分, 在内容上不能依赖大模型
(大模型的**局限**: 幻觉、正确性、时效性.....)

把握ChatGPT生成内容的特征

- 生成过程的敏感性:对提示词敏感
- 生成内容的不一致性:
 - 相同的提示,不同的上下文语境(相当于不同的提示)
 - 相同的提示,相同的上下文语境(Temperature参数)
- 注重生成内容的多样性
- 人工审核:不能100%相信大模型生成的内容

目录

一、什么是ChatGPT？

二、ChatGPT如何辅助学术写作？

三、案例分析

四、其他学术写作润色辅助工具

五、交流讨论

ChatGPT辅助学术写作的提示学习方法

- 润色 : Polish/Refine/Improve the following sentence/paragarph in an academic writing style:
[Your Text]
- 摘要 :
 - Refine the sentence for an academic NLP paper as the beginning of an introduction part: [Your Text]
 - Write a paragraph to first argue the importance of constituency parsing for general NLP tasks from a macro perspective, and then emphasize its importance for the field of computational linguistics / natural language processing for historical and ancient languages. [Your Text]

●

目录

一、什么是ChatGPT？

二、ChatGPT如何辅助学术写作？

三、案例分析

四、其他学术写作润色辅助工具

五、交流讨论

其他学术写作润色辅助工具

- Quilbot
 - <https://quillbot.com/>
- Grammarly
 - <https://www.grammarly.com/>

References

Kocoń, Jan, et al. "Chatgpt: Jack of all trades, master of none." arXiv preprint arXiv:2302.10724 (2023).

Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).

Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

Radford, Alec, et al. "Language models are unsupervised multitask learners." *OpenAI blog* 1.8 (2019): 9.

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

Ouyang, Long, et al. "Training language models to follow instructions with human feedback." *Advances in Neural Information Processing Systems* 35 (2022): 27730-27744.

目录

一、什么是ChatGPT？

二、ChatGPT如何辅助学术写作？

三、案例分析

四、其他学术写作润色辅助工具

五、交流讨论

Q & A

感谢聆听，欢迎交流