

LM 在计算机视觉领域的研究和应用

挑战和机遇

宫文娟

中国石油大学
青岛

APRIL 24, 2023



第一部分：LM 进展

I: 大规模语言模型 LLM

- 2020 年 6 月，首次发布 GPT-3 (Generative Pre-trained Transformer 3, ChatGPT core)，GPT-3 基于 Transformer。**Transformer** 的优点是通过堆叠效果会变好。
- 2022 年 11 月 30 日，OpenAI 公布了 ChatGPT。
- 2023 年 3 月 14 日，OpenAI 公布了 GPT-4。
- 2023 年 3 月 28 日，斯坦福发布了指令跟随模型 Alpaca¹，公布了 52K 的指令跟随数据集、生成数据集的代码、微调模型的代码、可以从公布的权重就计算 Alpaca-7B 模型权重的代码。
- 2023 年 3 月，加州大学伯克利分校、斯坦福大学和加州大学圣地亚哥分校联合研发了 Vicuna-13B 模型是利用从 ShareGPT 收集的用户共享的对话对 LLaMA 模型进行微调得到的。初步评估证明 Vicuna 可以达到 OpenAI ChatGPT 性能的 90%。

¹<https://crfm.stanford.edu/2023/03/13/alpaca.html>

I: 大规模语言模型 LLM

■ Vicuna 与以往模型效果对比:

Table 1. Comparison between several notable models

Model Name	LLaMA	Alpaca	Vicuna	Bard/ChatGPT
Dataset	Publicly available datasets (1T token)	Self-instruct from davinci-003 API (52K samples)	User-shared conversations (70K samples)	N/A
Training code	N/A	Available	Available	N/A
Evaluation metrics	Academic benchmark	Author evaluation	GPT-4 assessment	Mixed
Training cost (7B)	82K GPU-hours	\$500 (data) + \$100 (training)	\$140 (training)	N/A
Training cost (13B)	135K GPU-hours	N/A	\$300 (training)	N/A

I: 大规模语言模型 LLM

- 2023 年 4 月 3 日，伯克利 AI 研究院公布了 Koala²模型，也是在 LLaMA 模型基础上进行了微调。

Model	Training Set	Training Code	Public Weights	Dialogue Fine-tuned	Evaluation Method
Alpaca	OpenAI API outputs	✓	✓	✗	evaluation by 5 humans
ChatGPT	proprietary	✗	✗	✓	proprietary
Koala	Public dialogues & preferences	✓	✓	✓	evaluation by 100 humans

- 2023 年 2 月，Nebuly 公布了 ChatLLaMA 模型^{3,4}，提供了类似于 ChatGPT 的基于人类反馈的强化学习 (RLHF) 的架构来训练 LLaMA 模型。
- 2023 年 4 月，FreedomGPT 发布⁵，该模型基于 Alpaca。

²<https://bair.berkeley.edu/blog/2023/04/03/koala/>

³<https://www.nebuly.com/modules/chatllama>

⁴<https://github.com/nebuly-ai/nebullvm/tree/main/apps/accelerate/chatllama>

⁵<https://github.com/ohmplatform/FreedomGPT>

I: 大规模语言模型 LLM

- 2023 年 3 月，加州大学伯克利分校公布了 ColossalChat⁶。
- 2023 年 3 月，Databricks 公布了 Dolly。Dolly 是全世界第一个真正开源的指令微调的 LLM。

⁶<https://github.com/hpcaitech/ColossalAI/tree/main/applications/Chat>

I: 大规模语言模型 LLM-训练自己的模型

- 2023 年 4 月，prelit 发表的博客介绍了如何训练自己代码书写的模型⁷。主要针对自动写代码这个功能。

⁷<https://blog.replit.com/llm-training>

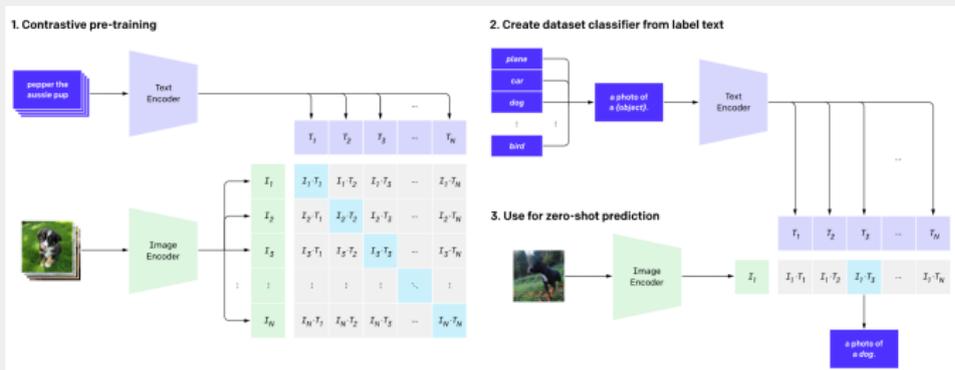
I: 大规模语言模型 LLM—低排减碳

- 人类大脑思考时大约是 22 瓦特的能量；GPT-3 模型释放 552 吨 CO₂，大约相当于旧金山到纽约来回飞 460 次。
- 2021 年 9 月，Meta [2] 提出了持续 AI 的概念。
- 2023 年 2 月，Meta 公布了模型 LLama-13b [3]，这个模型比 GPT-3 效果要好但是只需要 1/24 的碳排放。
- 2023 年 3 月 15 日，Pytorch 2.0 包含了降低总体训练时间、加速测试并降低训练和运行 AI 模型的碳排放量。
- **模型增长技术**：研究人员拷贝神经元（甚至是旧版本的所有 Transformer 层），然后堆叠起来，以增加模型大小。
- 2023 年 3 月，ICLR 的工作 [1] 利用已经训练过的小模型构建大模型，跟从零开始训练的大模型相比，这样可以节省 50% 的计算耗费。⁸。

⁸<https://vita-group.github.io/LiGO/>

II: 大规模多模态模型 LMM

- 2021 年 1 月 5 日, OpenAI 开源了 CLIP 模型。



- 2022 年 7 月 12 日, 首次公布 Midjourney 模型。
- 2023 年 4 月 5 日, MetaAI 公布了 SAM (Segment Anything Model) 模型。
- 2023 年 4 月, 威斯康星-麦迪森大学、微软研究院和哥伦比亚大学联合研究并公布了 LLaVA 模型⁹, 号称可以达到 GPT-4 的水平。

⁹<https://llava-vl.github.io/>

Appendices

REFERENCES



PEIHAO WANG, ET AL. *Learning to Grow Pretrained Models for Efficient Transformer Training*. ICLR, 2023.



CAROLE-JEAN WU, ET AL. *Sustainable AI: Environmental Implications, Challenges and Opportunities*. ICLR, 2023.



HUGO TOUVRON, ET AL. *LLaMA: Open and Efficient Foundation Language Models*. ARXIV.