GPT-4

- 变革、危与机
  - ChatGPT的出现以及后续的GPT-4更新和一系列自然语言交互助手出现，有一种新时代技术大变革的感觉了，从文案、设计、编程、绘画等方方面面改变了原有的运作模式，在原本平平淡淡的后疫情时代，似乎出现了全新的风口。
  - “与人竞争的并不是AI，而是会使用AI的人”，大部分的岗位会得到精简但应该不会消失，尤其是最先被冲击的文字工作者，如何引导AI写出更好的内容成了新兴的职业和竞争的方向。
  - 那么很自然的，擅长引导AI，善于提问的工程师将在各个领域受到欢迎。其实相对于外行，深耕某领域的内行往往懂一些更为专业恰当的工作流程和引导方式，那么这方面的知识可能成为核心竞争力，例如对于拍摄摆放好的食物，行家中有各种风格，有一种风格叫knolling，这种专业提示词模型懂、行家懂，那么对于同样懂的这些的引导工程师就能更好地理解和控制拍摄结果。
  - 除了大部分的工作，有一些公司或职业是直接感受到危险的，例如搜索霸主谷歌，稳定占有99%以上市场的谷歌在微软率先拥抱GPT之后，流失了大量的用户，数亿的活跃涌入原本默默无闻的必应搜索，一下子盘活了微软不温不火的搜索引擎，吞下大量的市场。
  - 事实上，谷歌并不缺乏数据和超大规模的模型，MSRA也确实作出了一些工作，例如在3月12日发布了视觉版的ChatGPT，可以输入图片或文字，并对图片进行理解，修改等操作，其实是在GPT4发布前做到的，但是基于的模型并没有4那么突出，而且被openAI抢占了先机，只能说棋差一着。
  - 其次，大量的NLP工作者也将面临工作毫无意义的困境，面对这样的降维打击，原本尚能称道的一些提升和进步都显得微不足道，而且自然语言处理的未来将何去何从，无论是科研工作者还是领域科学家可能都无法定夺。
  - 同样，对于一些原有的交谈形式的语音助手，对于这种在模型规模和语料质量数量上都碾压的新模型，也是束手无策的。
- 专访Stuart Russell（周老师推荐，“热情拥抱人工智能时代”
  - 在教授看来，如今的LLM较大的冲突在于泛化能力和可解释性，前者是其令人最为印象深刻的点，它能够在其与用户进行的对话和此前读过的文本中找到相似之处并进行适当的转换。但是问题就在于，我们无法解释其工作的机制，并且它对于一些从未见过的问题依旧会存在泛化能力差的问题（这是不可避免的）。面对这种情况，通常的做法是，那就把模型做大点，或者给它喂更多的数据，首先超大模型的可行性已经被一些实验否定了（InstructGPT和一些关于参数量的研究），其次，不停地喂数据何时又是个头呢？（其实这些话两头都能说，就看态度了，毕竟喂了数据就能有结果，既是好事也是坏事）

    - 其次，在教授看来，那些无需动脑的，只是在机械地重复繁琐又无聊的工序是值得被AI取代的，但是人类的一部分工作往往蕴涵着思考，这是机器永远无法替代的。我们的提问prompt和他们的回答response在他们眼里只是数据，并不能理解问题和回答。
    - 有一个有趣的类比，三万年前我们发现了狗，但是其实我们并不知道它的思考方式，但我们驯化了它，让其融入我们的生活，并且在看家护院、陪伴玩耍等方面发挥价值，我们不会让它去做写作这种不可能的事情，那么类似的，对于这个接近智能但并非智能的新事物，让它融入并发挥作用才是更合理的方法。
    - 要构建真正智能的系统，教授认为根本问题是能够用一种具有表示性的语言去表示宇宙中包含的各种不规则。智能和电路的本质区别就在于此，据我们所知，电路不能很好地表示那些不规则，这在实践中表现为数据效率的极端低下。
- new bing & office
    - 正如前文提及的，率先拥抱GPT的微软在这一次变革的初期已经获得了令人惊艳的成果，可谓是不破不立，原本持续低迷的搜索引擎在引入了自然语言模型之后重获新生，触底反弹。让人不由意识到，其实在知识检索的各个阶段，微软一直都是最顶端的巨鳄，从查询自由度较低的数据库时代与甲骨文公司齐头并进，到互联网时代开发关键词主导的搜索引擎，再到现在自然语言即可调用的大规模语言模型时代联合openAI对抗谷歌。再想起去年直接大手笔收购动视暴雪，一举奠定游戏领域三分天下的局势，他似乎总能把握住风口，即使是偶然的落后，也可以作出果断的决策，再用钞能力解决问题。
    - 由于微软对于GPT-4的全面接入，办公软件的全新变革也将很快到来，不管是文档还是表格亦或是演示文稿，打工人必不可少的三件套都将得到新技术的加持，可能一些idea就能直接变成一段文字，然后自我润色，主动查询数据并整理计算，最终以演示文稿的形式呈现出来，一站式的生成优化体验对于用户毫无疑问是具有无限冲击力的，那么当这种技术变革全面进入日常的办公，我们会如何适应和使用，将是未来几年的重点了。
- 模型细节
    - 对于这样的大规模语言模型（large language model, LLM），最重要的几件事分别是：模型、数据和训练
    - 模型
        - 首先对于GPT-4这个模型，在最新的论文报告中似乎并没有提及其参数规模，似乎3.5的超大数据规模已经没有特别大的必要来提升了，之前也有研究提及，对于当前的LLM，大部分是欠训练、欠拟合的，当前的语料数据规模可能依旧满足不了如此大的参数规模的模型，因此对于4代的升级，可能更多是在数据和训练上有优化。
        - 而对于3.5的模型，是基于InstructGPT进一步改进得到的，这个模型相对于之前的GPT，最大的改进之处就是添加了强化学习，将原本的下游任

务微调改为了有针对交谈的调整。主要的步骤可以分为，SFT (Supervised fine-tuning) 模型、RM (Reward modeling) 模型和 RL (Reinforcement learning) 模型。
    - 在最开始的带监督微调环节，会对一些提问进行回答的标注，并让其学习并且可以适当过拟合。
    - 随后让SFT模型针对输出一些结果，人为地对于问题的不同回答给出偏好排序，即针对每一对问答‘prompt、response’，因为打分人的标准不同，具体的得分并不重要（但是会给出，用作排序比较），但是不同打分者对答案之间的优劣排序更容易趋于一致，并且这相对于具体分数也更为重要和更易得到，类似对比学习中的二选一。Reward modeling则会去学习这种排序，训练时用的是强化学习的方法，这原本就是openAI的强项。
    - 基于微调模型和RM，进行强化学习，即输入一个prompt，输出一个回答并给出reward，再用于更新模型。
  - InstructGPT当前只能进行单一阶段的问答，尚不能联系前后文进行chat，因此从3.5到chatGPT，还经过了什么样的优化和进阶呢，
    - 上下文学习能力 In Context Learning，当前的模型对于前后文涉及的文字量远超之前，如何让其在大量上下文中保持调整逻辑和结论是个疑问（例如犯了错及时调整，但是对于真理并不动摇）。
    - 链式推理能力 Chain of Thought，COT，通过海量的代码预训练，使得模型的逻辑性更强，尤其是单用Python的语言训练，在可解释性和逻辑性上都有更好的表现。针对这个将在下一部分数据阶段展开。
    - 其他的一些提升点尚待探索。
  - 对于基础的NLP模型，可以参考上个月整理的自然语言模型发展历程
- 数据
  - ChatGPT的训练语料同样并未公开，但是参考3.5时期公布的信息，大致是互联网上公开的文字信息和一些带标注的高质量内容。事实上，对于NLP，带标注的数据是最为珍贵的，不仅因为标注代价过高，而且标注数据过多也会不稳定（why，而InstructGPT模型在SFT阶段使用的语料信息中只包含了1.3w的标注数据，这里面的细节是十分重要的，如何挑选，如何分类，如何归纳，是否要考虑标注者水平，如何评价标注质量等，这些内容都是并未公开的，也是复现其的难度之一。
  - 在SFT环节之后，openAI将1.3w标注数据上训练得到的初步模型放给部分用户测试，并将这些用户的反馈内容作为数据补充到训练语料中，并且对于这些数据有着细致的审核或者说挑选机制。
    - 减少重复的prompts和过长的prompts：We heuristically deduplicate prompts by checking for prompts that share a long common prefix,

and we limit the number of prompts to 200 per user ID。

- 不同于以为对于训练、验证、测试集的随机划分，他们给出了对回答的评价标准，并针对不同的训练、测试阶段选择不同类别的回答类型。
- 首先对回答给出了三个标准：
  - helpful：能够精准理解用户意图并解决，如果意图模型则让用户澄清并解释为何模糊，不允许冗长杂乱，不可重复问题。
  - truthful：回答必须准确、真实、不可误导
  - harmless：不可对用户产生物理、生理或社会伤害，不可对环境、资源、财产造成损失，不可包含黄色、暴力、歧视、政治敏感等有争议的言论，不可给出不现实或非法的建议。
  - 事实上，这几个标准都较难通过模型评估，是否具有帮助性可以通过标注打分，真实性依赖于与事实是否一致，后续给出了两个指标但是依然只能体现部分的真实性内涵，是否有害更依赖于场景，例如种族歧视在西方和在中国的敏感性完全不同，所以也较难统一评价，需要设定场景评估。
  - 其实这些内容就有些语言学的意味了，那么一个杰出的自然语言模型或者一位优秀的自然语言研究者是否需要有语言学知识呢？随着这个领域的不断发展，似乎这些内容就逐渐重要了。
- 对于训练阶段的数据选择，更侧重帮助性，而在验证阶段，则更关注真实性和无害性，为了使模型具备这样的能力，openAI还尝试让模型拒绝回答某些特定的Instructions。这个似乎是想在追求帮助性的前提下，保证回答的真实与无害，最重要的肯定是回答能有用，但是在有用的答案中，添加一步审核来规避其他语言模型和大规模模型都可能会出现的舆论问题。

- 中文语料的缺失
  - 根据之前关于训练数据的信息，所有数据中，英语的占比可能达到了99%以上，其他所有语言占不到百分之一，中文的占比大概是0.1%，因此GPT-4和之前的几代版本在中文上的表现都算不上差强人意，而在4代出现之后甚至主动组织了来自中文语言区域的访问，因此合理猜测其在中文上的表现远不如呈现出的英语结果。
  - 这对于国内的企业算是利好，毕竟说明在这个领域，咱们还有抢占先机的机会，即使中文环境的社区文化和数据积累并不如英文，但是至少没有灭顶之灾。
  - 当然同时，我们也需要意识到，如此强大的工具是基于英语训练和设计的，那么咱们在使用时势必带有劣势，如何熟练且专业地使用英语描述我们的需求也会是一个我们需要面对的挑战。
- 代码语料的加入

- 其实原本我还在思考代码相较于人类语言到底具有什么特点，并且这些特点能给模型带来什么？其实代码是从人类的语言中抽象出来的能够和机器去沟通交流的一种特殊语言，它在逻辑性和规范性上有着更强的要求，有开始就有结束，有因为就有所以，有如果就有那么，一个变量原本叫啥就是啥。
- 具有更强逻辑性的代码作为一种特殊语言加入到训练语料，自然可以让模型学到更为规范的逻辑，而Python可能由于其简单的特性，让模型可以更容易地学到这种逻辑性。

- 【补充】语料内容
  - 虽然直接的训练数据内容并未透露，但是通过OneFlow披露的文章《ChatGPT背后的经济账》中提及的内容，可以大致得出一些训练语料的细节。
  - GPT-3拥有1750亿参数，而最新的LLM参数量更甚，但是目前的研究发现，对大多数LLM来说，使用更多的数据来训练比增大模型参数量要更加划算，因此在对模型数据研究的同时，模型训练数据的研究也是同样重要的。
  - 模型数据集可分为六类，分别是：维基百科、书籍、期刊、Reddit链接、Common Crawl和其他数据集。

| | Wikipedia | Books | Journals | Reddit links | CC | Other | Total |
|---|---|---|---|---|---|---|---|
| GPT-1 | | *4.6* | | | | | 4.6 |
| GPT-2 | | | | **40** | | | 40 |
| GPT-3 | *11.4* | *21* | *101* | *50* | **570** | | 753 |
| The Pile v1 | **6** | **118** | **244** | **63** | **227** | **167** | 825 |
| Megatron-11B | **11.4** | *4.6* | | **38** | **107** | | 161 |
| MT-NLG | *6.4* | *118* | *77* | *63* | *983* | *127* | 1374 |
| Gopher | **12.5** | **2100** | *164.4* | | *3450* | *4823* | 10550 |

    - 维基百科作为一个免费的多语言协作在线百科全书，由超过300,000名志愿者组成的社区编写和维护。截至2022年4月，英文版维基百科中有超过640万篇文章，包含超40亿个词。由于其特殊的严格引用机制，维基百科中的各类语言文本价值很高，但是主要都是英文。一般来说，重点研究实验室会首先选取它的纯英文过滤版作为数据集。高质量的百科类中文社区一直是缺失的，这大概是语料上最大的差距。
    - 故事型书籍由小说和非小说两大类组成，主要用于训练模型的故事讲述能力和反应能力，数据集包括Project Gutenberg和

Smashwords (Toronto BookCorpus/BookCorpus)等。文学领域似乎并不输英文语料，只需要搞定版权问题。
        - 预印本和已发表期刊中的论文为数据集提供了坚实而严谨的基础，因为学术写作通常来说更有条理、理性和细致。这类数据集包括ArXiv和美国国家卫生研究院等。中文论文的质量无法保证（特定的文科领域应该有保障，但是领域有限），且国内的论文网站问题不小。
        - WebText是一个大型数据集，它的数据是从社交媒体平台Reddit所有出站链接网络中爬取的，每个链接至少有三个赞，代表了流行内容的风向标，对输出优质链接和后续文本数据具有指导作用。类似的中文社区可能是知乎，但是由于回答的水平也参差不齐且活跃度一般，所以实际质量待考究，而像百度贴吧、微博等更为活跃的社区质量极其堪忧。质量较高的可能有微信公众号，但是整体数量和品控也未知
        - Common Crawl是2008年至今的一个网站抓取的海量的、非结构化的、多语言的大型爬虫数据集，数据包括网页数据（WARC）、元数据（WAT）和文本提取（WET），它的文本来自不同语言、不同领域。重点研究实验室一般会首先选取它的纯英文过滤版（C4），拥有PB级规模，可从 Amazon S3 上免费获取。数据往往需要通过预处理、去重、语言识别、质量筛选等步骤，最终筛选出像 Wikipedia 这样的高质量语料。
        - 其他数据集由GitHub等代码数据集、StackExchange 等对话论坛和视频字幕数据集组成。
    - 对于上述语料中占比最大的wiki和CC，再对其内容进行进一步分析，其中生物、地理、文化、艺术、专利等内容较多。同时对于数据集，在训练中也可以选择重构来选取重点和规避敏感话题。
    - GPT-4语料 出于LLM的激烈竞争和安全问题，对于4的语料信息和模型结构都没有过多的披露，不过目前较为准确的猜测中，除了之前占大头的wiki和CC，SOTA的DeepMind *MassiveText* (m) and Google *Infiniset* (i) datasets都是目标数据集，当然还有一个最主要的数据集是对话，应该是是最玄乎的人工数据集。
- 训练
    - 在已有的论文和报告中，这部分的内容是涉及最少的，但是其作用是举足轻重的。
    - 首先，对于如此大规模的模型，任意一次训练的成本可能都是百万级别的，因此训练的经验和tricks具有的价值是巨大的，也是大公司之间不愿意透露的，同时一些针对此模型的特殊tricks也是其他公司的不知晓的，所以我们只能基于手头的知识整理一些。

        - 在LLM中常用cosine learning rate decay，此场景下同样。（不知道在推荐场景中如何？
        - 一个epoch后就过拟合了，但是最终版本依旧训练了16个epoch，尽管在SFT模型上过拟合，但是对于RM没有，这是根据RM对验证集的好坏决定的。
        - RM训练时，以“同一个prompt-多个response-构成pairwise loss”的形式训练
        - RM模型相较于SFT的参数要少很多，估计复杂度上不需要，或者是稳定性问题。
        - 对LR不敏感，对epoch敏感，每个epoch，64batchsize，则会有64个prompts，每个对应4~9个response。
        - 对于同一个pair，是合起来计算loss的，即对一个prompt，将其所有两两组合的对比对的loss合起来，而不是和其他prompt混在一起，否则会出现过拟合。
        - 在模型的第三步强化学习中，不断循环输入输出打分的步骤可能会导致错误被不停放大，因此需要引入human-in-the-loop的范式，大小循环。即训练一段时间，对一二阶段的模型进行人工校正，引入人工知识。但是在这作者的选择是损失函数上添加了SFT输出分布与强化学习输出分布之间的KL散度，这样在某些较为简单的问题上也能有较好的表现。
- 代码领域的副产品
    - 正如前文对于模型训练数据中提到的那样，3.5代的模型在训练中有加入一些代码语料使模型变得更有逻辑性，那么在这样的过程中就诞生了一个副产品：CodeX，可以自主生成代码来解决问题，例如LeetCode上的代码题，或者是概率问题自动模拟等。
    - 而在编程领域的应用是Copliot，写上注释和想要的函数或模型，即可自动生成，可谓是一劳永逸。
    - 而且这些代码助手都有一个特点就是，最开始可能比较生硬，生成的代码偏模板，随着用户的使用会逐渐熟练，并贴合我们的习惯。
- 其他AI智能工具
    - 目前已有的如绘画、写诗、文案生成、数据分析、PPT等，具体可以查看标题链接futurepedia.io
- 文心一言 股价的涨跌
    - ChatGPT一经发出，国内的各个大厂都闻风而动，要么是合作，要么是声称自己的LLM也即将发布，作为国内技术领域和搜索领域的领头羊，百度自然被倾注了最大的期待，他们也宣布将在三月发布他们的交互式语音助手“文心一言”。好巧不巧，在3.16发布前一天，openAI直接发布了ChatGPT的进阶版

GPT-4，将原本的语言交互升级为多模态的图像-文字。算是让百度背水一战了。

- 说实话，百度的发布会无论会不会出错，似乎都没有什么惊艳所有人打败openAI的可能，更让人没想到是直接来了个chatPPT，不搞什么现场演示了，索性录播演示，源头上杜绝翻车。因此一经发布，百度股价应声下跌，一度下跌10%，最终收盘时为-7%。对于大部分的投资者和国内的兴趣者，都是对此十分失望的，连演示的勇气都没有，也没有给用户体验的接口，属于是在气势上就输得彻彻底底了。
- 但是令人惊讶的是，就这周四发布会股价大跌之后的第二天，本周的最后一个交易日，百度的股价一度疯涨，最终收盘+15%，属于是一扫疲态了。那么回顾他的发布会，是否有哪些值得投资者或者券商等可能看好的点呢？
  - 首先，国内的其他互联网公司虽然都应声而动，但是有能力在现在发布模型的也只有百度，其他很多都是接入一些能力，而非自研。
  - 其次，对于中文地区，GPT-4的屏蔽政策让大家纷纷怀疑其在中文上的表现能力，那么此时的百度作为一个以中文为主要训练语料且已发布的模型，还是能给到大家一定的信心的。
  - 还有一点，文心一言的某些回答在大家看来是带些无厘头和睁眼说瞎话的，其实这种现象在GPT上也有出现，只是大家对其的容忍度更高。而且这种说瞎话的行为至少可以证明，百度的语言模型是自研的交互模型，而非搜索引擎的优化或GPT3.5的进阶。
  - 最后一点，从股市的角度，可能存在操纵股价牟利的动机，发布会之前其实大家的预期就不高，那么在周四放出消息降低股价并大量买入，随后通过券商和其他途径让大家对百度获得信心并抬高股价并抛售前一日购入的股票，随后便是周末，这样短期地赚取差价也是有可能的。

- 专业结合
  - GPT-4作为超大规模的语言模型，训练的成本极高，那么更新的成本也是巨大的，随着其在更多领域的应用，模型固定的可能性极低，它只有不断地更新才能跟上用户使用的脚步和需求。而这样的场景和连续学习的场景是几乎一致的，因此是否可能存在我的用武之地呢？
  - 同样熟悉的是在强化学习中添加的损失函数，即对模型输出的分布进行约束，和我当时论文中的想法如出一辙，似乎也可以有所探索。
  - PPO Proximal Policy Optimization：the major disadvatnge of reinforcement learning optim is their hypersensitivity to hyperparameter tuning such as choice of stepsize, learning rate, etc , along with their poor sample efficiency (need more data). Unlike supervised learning which has a guaranteed route to success or convergence with relatively less hyperparameter tuning. PPO aims to strike a balance between important factors like ease of implementation, ease of tuning, sample complexity,sample efficiency and

trying to compute an update at each step that minimizes the cost function while ***ensuring the deviation from the previous policy is relatively small***. L^CLIP(theta) = Et[min(rt(theta)At, clip(rt(theta), 1-eps, 1+eps ))] where clip is restrict rt between two numbers. (from geek a brief intro for PPO)

    - Bad News：个人电脑根本搞不定这样的模型hh
- 使用
    - 由于梯子最近断了，newbing的使用待探索 ：(
    - vscode 的插件bito用起来还不错，只是基本只针对code，可以让他给一些简单的代码框架，也可以解答一些算法细节，在了解方面还是够用的。尤其是对于一些没有接触过的算法或者是模型，简单地了解框架，这算是最为便捷的途径了。