



ChatGPT 为何没能诞生在中国？

上接 06 版

习等)已进入到一个瓶颈阶段。”李笛告诉《中国新闻周刊》。

刘知远介绍, 预训练模型为自然语言处理带来了两个变化: 一是可以充分利用网上海量的未标注数据, 模型的规模和能力得到显著提高, 因而, 从规模角度, 预训练模型被称为大模型; 另一个变化是, 大模型具有非常强的通用能力, 只需经过少量参数微调, 就可以用于机器翻译、人机对话等不同任务。“大模型思想, 某种意义上有点像‘大力出奇迹’, 把大量数据压到一个很大的黑盒子中再提出来。”李笛对《中国新闻周刊》说。

“但人工标注还是一个重要的数据来源, 此前大量的标注没必要了, 但在特定任务上还需要标注一些, 比如希望这个模型输出的内容更符合人的需求。”刘知远说。近期, 《时代》周刊的调查, 呈现出 ChatGPT 智能背后的灰暗角落。《时代》周刊称, 2021 年 11 月, 为训练 ChatGPT, OpenAI 使用了每小时收入 1~2 美元的肯尼亚外包劳工, 对性虐待、仇恨言论和暴力等文本进行标注, 保证聊天机器人过滤有害信息, 输出适合日常对话的内容, 同时, 这对标注员的精神和心理产生极大伤害。

学术界对大模型态度一直存在分歧。北京智源人工智能研究院副院长刘江介绍, GPT-3 论文发布时, 无论国内外, 不少自然语言处理领域学者认为, 大模型只是靠持续烧钱、粗暴扩大数据规模提升能力, 并非真正创新。刘知远对《中国新闻周刊》说, 还有一个更现实的问题, 大模型需要极大数据和算力支持, 如果一些研究者的实验室没有算力支持, 他们可能会选择过去熟悉的方向。

OpenAI 是全球所有科技公司中, 大模型的坚定支持者。2019 年, OpenAI 推出参数为 15 亿的 GPT-2, 2020 年推出 GPT-3, 将参数提升到了 1750 亿, 成为当时全球最大的预训练模型, 引发

业内轰动。“绝大部分人根本没想到, 人类可以把一个模型训练到这么大的规模, 这对自然语言交互的流畅性有非常强的提升。”刘知远说。

参数增多, 使语言模型学习进阶到更复杂模式。早在 2020 年, GPT-3 可以做到其他模型无法做到的事情, 比如作诗、写复杂的文章和代码等, 通用性极强。刘知远形容, GPT-3 像是一个伶牙俐齿的人, 有不错的表达能力, 但不具备很强理解能力。

2022 年, GPT-3 进一步升级为 GPT-3.5, 这是 ChatGPT 的底层基础, OpenAI 进行微调, 提升交互能力, 让它“听得懂人类的复杂指令”。“这些都经过了专门训练, 像父母对小孩的调教。”刘江形容, GPT-3 像是两三岁的天才儿童, 读完世界上所有的书, 但不知轻重, 也没有价值观, 需要父母精心教育和启发, 让它在聊天方面发挥潜力。

清华大学智能产业研究院首席研究员聂再清向《中国新闻周刊》介绍, ChatGPT 聊天能力的显著提升, 是引入了一个新的数据训练方法, 人类反馈强化学习(RLHF)。OpenAI 引入人类评判员, 创建一个奖励模型——评判员不断地跟 ChatGPT 对话, 并对它生成的答案按照质量好坏评分, 模型收到反馈后进行优化。山姆·奥特曼也承认, “让模型以特定方式对人们有所用途, 并找出正确的交互范式, 却得到了惊人的效果。”

在自然语言处理领域, 流传最广的一句话来自于比尔·盖茨: “自然语言处理是人工智能皇冠上的明珠, 如果我们能够推进自然语言处理, 就可以再造一个微软。”

刘知远认为, ChatGPT 推出后最大的价值在于, 能用 RLHF 等技术, 把大模型的能力展现出来, 让公众意识到, 人工智能与人类自然语言的交互达到非常高的水平, 机器已经可以“能言善辩”了。

但大模型为何有时生成错误答案? 在李笛看来, 这是由其技

术结构决定的。聂再清进一步向《中国新闻周刊》解释, 这是因为 ChatGPT 本质上还是一个基于概率的语言模型, 本身不涉及知识的对错, 未来仍需用更多的专业知识训练。

当下, 大模型与搜索引擎结合已是大势所趋。聂再清建议, 新版搜索引擎给出综合答案后, 最好附上原始网络链接, 有利于用户自己验证 AI 回答的正确性。目前新版必应在每个回复下加入信息来源。但两者结合最终成功与否的关键, “还是在于 AI 总结的答案绝大部分都是对的, 不会耽误用户花更多时间来验证结果。”

更关键问题是, ChatGPT 及其背后的大模型, 仍是基于数据驱动生成内容, 不是像人类一样会思考和推理。但 2 月上旬, 斯坦福大学计算心理学教授迈克尔·科辛斯基发表论文称, 在对几个语言模型进行专业测试后, 他发现 ChatGPT 表现接近 9 岁儿童的心智。

近期, 因在深度学习领域贡献获 2018 年图灵奖的杨立昆谈到, 人类思维方式和对世界的感知, 是人类获得常识的基础, 聊天机器人的模型没有这种能力。对此, ChatGPT 回答《中国新闻周刊》说, “我的设计是基于概率模型和大量的数据训练, 以回答问题和执行任务。我不具有意识、情感或主观体验, 也不能对世界产生真正的影响。”

随着 ChatGPT 的发展, 未来会不会替代人类的工作?

“我相信 ChatGPT 会取代一些工种, 或者让一些工种不需要太多人参与, 这是一个潜移默化的过程。”刘知远对《中国新闻周刊》说, 但与此同时, 它也会催生一些新的工作, 比如, 以前画画需要很高的门槛, 但现在, 即使一些人不会画画, 但有天马行空的想象力和创意, 一样可以和 AI 一起创作。

ChatGPT 似乎承认自己可以在部分工作中替代一些人力劳动, “可以在许多行业中使用, 例如客服、教育、媒体、医疗

保健和金融”, 但它补充说, “我不能取代需要人类情感和社交技能的工作, 例如教育和医疗保健等需要人类情感互动和洞察力的领域。”

中国何时会有自己的 ChatGPT?

相较国外, ChatGPT 在国内的热度稍显滞后。谷歌和微软短兵相接时, 国内搜索巨头百度也宣布 3 月将推出中国版的 ChatGPT “文心一言”。腾讯称, 在 ChatGPT 和 AIGC 相关方向已有布局, 阿里达摩院正在研发的类 ChatGPT 的对话机器人, 目前已开放给公司内员工测试。此外, 快手、京东、360 等多家互联网企业也都表示在相关领域研发和布局。

2 月 13 日, 北京市经济和信息化局在北京人工智能产业创新发展大会上明确表示, 北京将支持头部企业打造对标 ChatGPT 的大模型。

李笛提到, 在 ChatGPT 之前, 国内和国外已经有很多公司在利用大模型做很多产品和研发, 市面上也有很多训练出来的大模型, “只不过在人工智能的训练过程中, 研发者的专注度、投入度不一样”, 并不存在“技术壁垒”。ChatGPT 火热背后, 是 OpenAI 从 2018 年以来持续投入完善大模型, 取得了这一效果, 所以有一定“时间壁垒”。

2 月 7 日, 360 在互动平台表示, 公司人工智能研究院从 2020 年起, 一直在包括类 ChatGPT 技术在内的 AIGC 技术上有持续性投入, 但截至目前仅作为内部业务自用生产力工具使用, 且投资规模及技术水平与当前 ChatGPT 3 比还有较大差距, 各项技术指标只能做到略强于 ChatGPT 2。

早在 2020 年, 北京智源研究院曾推出超大规模智能模型“悟道”项目, 阿里达摩院自研预训练模型框架 ALICE。2021 年, 深圳鹏城实验室为首的联合团队, 推出参数为 2000 亿的大模型“鹏程·盘古”, 探索通用人工智能。多位受访专家提到, 中国目前大模型研发与 OpenAI 仍有差距, 国内要有像 GPT3.5 这样的大模型, 但没必要每个公司都去投入和研发。

算力离不开芯片。2 月 12 日, 国盛证券估算, 今年 1 月, 平均每天约有 1300 万独立访客使用 ChatGPT, 对应芯片需求为 3 万多片英伟达 A100GPU, 初始投入成本约 8 亿美元, 每日电费 5 万美元左右。而 GPT-3 训练一次, 成本约为 140 万美元, 对一些更大的大模型, 训练成本介于 200 万美元至 1200 万美元之间。这一成本对全球科技大企业而言, 尚在可接受范围内, 但并不便宜。

聂再清分析说, 中文很多高质量信息在 APP 里, “有点数据孤岛的意思”, 公开的高质量互联网语料可能不如英文多。另一个挑战是, 语料筛选、清洗、预处理和标注需要相关技术人员深度参与, 会有一个不断迭代和较为长期的过程。

此外, 中文机器语言学习在很多方面要比英文更复杂, 中文和英文在句法结构、缩写规范方面也有差别。聂再清提醒, 打造对标 ChatGPT 的大模型并非一蹴而就, 需要时间。

但李笛认为, 最终还是要看它能否实现“端到端”的落地。在国内, 绝大部分大模型都还无法实现这一目标。同样, 很多 AI 绘画单幅质量已很好, 但在可控性上却“漏洞百出”。所以今天大模型的应用普遍还停留在试用阶段, 距离真正大规模商用, 还有很多事情要调整。