

复旦团队发布 MOSS 模型,研究成果将开源

由美国 OpenAI 公司开发的 ChatGPT 模型在全球点燃了新一轮 AI 革命,海内外关注度陡增。前不久,复旦大学计算机科学技术学院邱锡鹏团队发布类 ChatGPT 模型 MOSS,新闻迅速登上微博热搜。

MOSS“火”了。2月20日发布当天就收到大量内测申请、采访、投资、合作邀约——公众对于这项原本局限于 NLP(自然语言处理)学术圈的新技术热情高涨。邱锡鹏团队对此颇感意外,但又很快回归平静。

大家最兴奋的一天,是春节前的腊月二十八。项目主开发者、计算机科学技术学院博士研究生孙天祥,在测试过程中输入了一个中文问题,MOSS 却以英文正确回答,“就像一个不会说但听得懂中文的人”。而当时版本的 MOSS 还很初级,中文语料占所有训练数据不到 0.1%。

“很神奇,我们没有教过它机器翻译。”MOSS 显示出的潜能让邱锡鹏当晚激动到失眠。他把 MOSS 比作一个“聪明的小孩”,即便现在还不擅长写诗、解题或很多具体的事,但已展示出成为通用人工智能 (AGI) 大框架的潜能,“很多遥不可及的事情,它一点就通了。”实际上,邱锡鹏也让 6 岁的女儿和 MOSS 聊天,发现孩子可以愉快地和 MOSS 对话很长时间。

MOSS 受到关注的背后,是科研人员十年如一日的积淀。作为一名人工智能研究专家,邱锡鹏从读博期间就开始涉猎机器学习,离校工作后进入自然语言处理研究领域。他和团队在自然语言处理的基础模型和基础算法上形成了很多创新的研究成果。邱锡鹏的著作《神经网络与深度学习》被广大读者亲切称为“蒲公英书”,在许多“人工智能必看书单”中榜上有名。去年,他还带领团队获得了中国中文信息学会“钱伟长中文信息处理科学技术奖”一等奖。

这些天,邱锡鹏与他的 MOSS 团队——8 位年轻的复旦学生——继续紧锣密鼓地开展内测和迭代工作。新模型预计在 3 月份底优化完成,后期再逐步对社会开放。

目前已参与内测的一些用户表示,尽管 MOSS 在参数规模上和 ChatGPT 相比小一个量级,事实性问题覆盖不够全面,经常会“一本正经地胡说八道”,但确实有“ChatGPT 那味儿”,“基本功能都实现了”。

邱锡鹏很乐观,认为在不远的将来,MOSS 这类大型语言模型会成为和搜索引擎一样常规的存在,为人们生活的方方面面提供助益。

开源(即开放源代码和模型参数)是邱锡鹏和复旦自然语言处理实验室的同仁们一直以来的学术坚持。“这次也会将研究成果开放给公众与社会。”他说。

利用有限资源做出类 ChatGPT 模型

问:团队 2 月 20 日发布了 MOSS 模型,是刚刚建成吗? 前后花了多长时间?

邱锡鹏:事实上,我们在春节前就开发出了第一代模型。它显示出了很大的潜能,与之前的聊天系统大不相同,有着不错的人类意图理解能力,也有很多涌现能力,比如未经训练就学会了机器翻译。之后,我们又花了一个多月的时间打磨它的工程部署情况,比如提高效率、优化界面等。

MOSS 的开发不是一蹴而就,它离不开我们团队过去的铺垫工作和长期积累的研究经验。从 2021 年起,我们就开始做中文生成式预训练模型,也开源供别人下载,每月平均有上万次下载。后面我们又提出“语言模型即服务”的概念,认为基础语言模型会成为语言服务的基座。2022 年,因为意识到大型语言模型会成为将来的基座,所以开始做大型语言模型方面的训练。后来又花费了半年时间,研究如何使大型语言模型理解人类指令以及具备对话能力。

问:MOSS 如何能够实现“端到端”走通大语言模型,克服了哪些难点?

邱锡鹏:“端到端”是一个学术概念,指的是从零开始,信息收集、数据处理、建立模型,到最后形成一个具有和人类对话能力的大模型,中间所有技术路径可以走通,这个就叫由起点到终点的“端到端”。因为 OpenAI 至今没有公布开发 ChatGPT 的技术路线和技术细节,所以我们需要靠有限的公开信息来自己摸索。

这个过程非常难,包含了非常多经验性、直觉性的设计,关键要打通两步:第一是基座,大型语言模型的基座不是简单的参数足够大就可以,还需要赋予大型语言模型各种各样的知识能力、学习能力,还有逻辑推理能力。第二就是要通过一些指令触发它的对话能力,让它理解人类意图,与人类能够交互对话。

到目前为止,我们还能把控制技术路线,但未来可能会面临更大的困难,因为我们收集了非常多和人类交互的指令,要赋予它价值观和各种各样的能力,就要请一些专业人士来帮助我们设计,进一步增强 MOSS 各方面的能力。

问:你们的团队构成如何?

邱锡鹏:我们自然语言处理实验室是上海市智能信息处理重点实验室的一部分,团队中老师和学生加起来近百人,一直积极贯彻学校和学院推进的有组织科研模式,聚焦自然语言处理和人工智能的国际前沿技术研究。和大型语言模型基座研究相关的学生有 30 余人。

自然语言处理领域最前沿的探索

问:MOSS 模型一经发布就备受关注,对此,你们有预想到吗?



▲ MOSS 团队合影(左六为邱锡鹏)

邱锡鹏:上热搜完全出乎我们的预料。关注度这么高,我想可能是大家对国内团队做出类 ChatGPT 模型比较兴奋,兴奋的原因则在于之前有很多声音,说我们和国外的技术水平差距非常大,想追上的话,要花很长时间。但我们的努力证明,其实不需要那么久。

问:国内外已有不少公司正在研发类 ChatGPT 模型,投入不菲。作为一个高校学术研究团队,你们开发 MOSS 模型的初衷是什么?

邱锡鹏:我们开发 MOSS 模型,是想在百亿规模参数上探索和验证 ChatGPT 的技术路线,证明我们在技术实现上并不落后于国外。另外也是想证明,这种技术也并非一定被大公司垄断,我们这样一个学术研究的实验室,在相对有限的资源下,也能够做出类 ChatGPT 模型。

当然,我们开发 MOSS 模型,不止于追求和 ChatGPT 能力相近的模型。MOSS 的定位是自然语言处理领域乃至是通用人工智能领域最前沿的探索。也许工业界更看重它的落地性能,而我们更看重它的下一代发展,即如何实现通用人工智能。从学术角度出发,只有我们比 OpenAI 看得更远、更超前,才可能最终实现超越。我们坚信,学术界在这个领域大有可为。

问:MOSS 的名字是怎么来的?

邱锡鹏:在学术圈,大家非常喜欢用影视形象给自己开发的 AI 模型命名,也是一个比较常见的做法。我们开发出这个具有对话能力的大型语言模型之后,也想找一个国产的、能够代表中国特色的影视形象来命名。

开发过程中,正逢《流浪地球 2》电影热映,我们的团队成员都非常喜欢《流浪地球 2》,也都是《流浪地球 2》的粉丝。影片中的智能量子计算机 MOSS 给我们留下了深刻的印象,所以我们就把模型命名为 MOSS,也是向《流浪地球 2》这部电影致敬。这几天也有《流浪地球 2》

的粉丝们给我们发邮件,希望我们一定要加油,真的能够做出来。

关注模型的逻辑思维能力和

问:MOSS 与 ChatGPT 相比,有哪些主要差异?

邱锡鹏:最大差异还是参数规模,另外一大差异就是迭代能力。参数量和交互数据量的差异,进一步导致了 MOSS 和 ChatGPT 在事实类知识储备上的差距,表现为 MOSS 在回答事实类问题时更容易出错。事实类知识遵从二八法则,80% 都是长尾知识,不知道就是不知道,但不代表模型不会学习。我们很容易通过扩大知识库,扩大模型的学习语料和参数规模,让模型具有更多的知识。因此,我认为,大家在关注 ChatGPT、MOSS 这种模型的时候,更应该看重它的理解能力、学习能力以及思维能力,而不是看重它的事实类知识储备。如果从逻辑类能力来看,我觉得 MOSS 的表现其实还不错。

问:MOSS 的科技创新点在哪里?

邱锡鹏:MOSS 的特点是小规模,比较容易适应个性化模型。从国内来看,大部分企业都有非常大的需求去接入 AI 服务,但如果直接把 ChatGPT 这么大的模型迁移过来,企业是用不起来的。MOSS 这种百亿级规模就非常合适,可以企业内部私有部署,经过一些数据微调就可以转化为生产力。

所以我们觉得,在 MOSS 这个级别模型上,可以赋予更多专业化能力,比如接入外部知识库,增加搜索或特定领域工具的能力。就像人类一样,很多能力也不够,但我们可以使用工具来把我们的能力放大,比如记忆不够,可以通过查词典、使用搜索引擎。对 MOSS 也是一样。如果它本身的知识量不是那么大,我们就要想其他方法,让它更善于使用各种各样的工具,从而为各行各业赋能。这个可能是我们未来会和 ChatGPT 最主要的差异。

通用人工智能未来可期

问:您认为中国版 ChatGPT 模型建设的必要性和价值在哪里?

邱锡鹏:首先,从大的方向来看,像 ChatGPT 模型,它并不向中国内地开放。中国要想站在大型语言模型或者将来的通用人工智能等技术的最前沿,就必须建设自己的语言模型基座。其次,国外开发者不太可能以中文为主去发展他们的模型,他们的发展重点还是在英文上。那么,我们要想开发一个大型语言模型的基座,用于国内的信息处理,特别是中文信息处理,就必须建设一个中文能力非常强的大型语言模型。

问:MOSS 的优化会包括哪些方面? 团队对 MOSS 模型的近期目标和最终期待分别是什么?

邱锡鹏:未来,MOSS 的优化将会围绕三方面展开。第一,我们会准备更高质量的中文数据;第二,我们会开放接口,让 MOSS 与人类进行对话,收集更多的对话数据;第三,我们会进一步加大投入,扩大它的参数规模,如果 MOSS 的参数规模能够上升到 500 亿或者 1000 亿,它的能力就又会大幅提升。

近期的目标,是希望 MOSS 成为国内技术比较领先的对话式语言模型。我们希望能坚持做一个不为利润所驱使的研究机构,把研究成果无偿地分享给学术界,也在合法合规的前提下把研究成果开放给业界,让他们去做定制化或者在特殊领域的应用。下一步进行顺利的话,我们会于 3 月底左右开源。

长远来看,我们期待把 MOSS 作为一个通向通用人工智能的基座,让它变成一个像科幻片角色的真实存在。我很乐观,我觉得通用人工智能由科幻步入现实,应该不会很远,也许 5 到 10 年。那时,我们会像现在接受搜索引擎一样,接受通用人工智能。

文 / 殷梦昊、许文婧
摄 / 戚心茹