

视觉中国供图



创新基础大模型 推动人工智能技术专业化应用

科技日报讯(记者龙跃梅)5月18日, AI赋能数字中国产业论坛暨2023云从科技人协同发布会举行。会上,云从科技集团股份有限公司(以下简称云从科技)展示了最新的人机协同操作系统,并宣布云从科技“从容大模型”的正式亮相。

ChatGPT等AI大模型的发布,引发了全行业对于AI技术的关注, AI行业迎来了新一轮发展浪潮。

此次发布会上,云从科技分析了AI大模型带来的内容和交互方式变革而引发的人工智能市场爆发,特别是与行业创新应用相互叠加产生的持续增长潜力。

云从科技董事长周曦认为,金融、法律、医学等不同行业,都有自身的行业大模型。不过,如果没有强大的基础大模型,直接去打造行业大模型,很难保持长期的生命力。这是因为,如果想让行业大模型足够实用,便需要重新训练基础大模型。此外,要想在产业实践中实现规模化应用,就必须严格控制行业大模型的成本,并提升其效率,这就需要掌握基础大模型。

据了解,云从科技经过多年积累,训练出足够强大的基础大模型——从容大模型。通过实时学习并同步反馈结果,从容大模型可以解决AI应用的痛点问题。从容大模型具备上下文学习能力,实现更好的交互性,特别在金融、游戏等应用场景,多轮对话技术在人机协同操作系统中将得到更充分的应用。云从科技现场演示了从容大模型的基础功能,包括趣味问答、中英文翻译、编程与阅读理解等。

会上,广州市政协副主席、市科技局局长王桂林表示,广州人工智能、大数据、云计算等优势产业不断集聚,打造了以人工智能与数字经济试验区为龙头的“一核牵引、多点支撑”空间发展格局,人工智能和数字经济产业规模和质量暂居国家第一方阵。

未来,广州将加速汇集人工智能与数字经济产业,持续培育壮大产业集群,推动创新链、产业链、人才链、资金链深度融合,持续提升产业发展核心能力。

建设智慧城市 需人工智能与多项前沿技术融合

◎本报记者 叶青

智慧城市建设是优化城市管理,促进城市治理体系和治理能力现代化,推进治理模式转型升级的重要路径和抓手。5月18日,2023自主可控GIS技术创新与应用发展论坛在深圳举行,会上多位院士专家围绕智慧城市建设进行了深入探讨。

“智慧城市建设需要超越部门的全局性思维、系统性思维和开放性思维,从纵向耦合变为横向耦合,通过泛在感知网络、传输网络、统一的大数据平台来支撑个性化应用。”中国工程院院士郭仁忠认为。

“城市是像人体一样的运行体、有机体、智能体,需要具备实时感知、动态监测等多种能力。”国际欧亚科学院院士邵伦认为,新一代的智慧城市建设需要融合人工智能、社会感知、物联感知等方法,创新时空智能耦合与城市数字孪生技术,构建时空智能体,实现行业场景的深度融合,使城市具备泛在感知与实时监测的能力、虚实结合双向映射与数据挖掘知识发现的能力、全生命周期模拟仿真的能力以及综合态势研判与智慧管理决策的能力。

面向数字中国、智慧社会建设需求和广东省高质量发展任务,国际欧亚科学院院士张新长科研团队系统研制了融合智能感知、智能融合、智能认知、智能决策、智能服务等城市的时空智能化信息平台。“让机器打破多元信息壁垒,理解城市时空演变过程,挖掘其背后隐含的本质,在人工智能与地理信息科学领域均属于前沿难题。”张新长表示,城市时空智能化信息平台或能为智慧城市建设与运行提供基础性支撑。

图说智能

博物馆里体验AI交互



5月18日,天津数字艺术博物馆正式向公众开放。作为该数字艺术博物馆的首个展览,“梵高的世界”全景数字艺术互动大展以梵高一生中所创作的多幅艺术作品为素材基底,运用AI互动、体感交互、AR(增强现实)等数字化技术手段进行展示。该展览在馆内设置8个沉浸互动区域,为观众带来视听融合的交互体验,使艺术、科技、观众深度互融,助力数字文化产业发展。

图为观众在天津数字艺术博物馆“梵高的世界”全景数字艺术互动大展展上体验AI交互。

新华社记者 孙凡越摄

中文大模型让AI更“接地气”

◎本报记者 陈曦

“请讯飞星火认知大模型模仿梁晓声先生笔下的小说《人世间》的风格,续写一小段文章。”5月20日,在第七届世界智能大会闭幕式上,主持人蒋昌建向讯飞星火认知大模型发问。短短几秒钟,续写文字便“跃然纸上”。原作者梁晓声认为,续写内容简练,文字有一定温度,从传达的情感和思想来看几乎“无可挑剔”。

在本次世界智能大会上,生成式人工智能毫无疑问成为大家关注的焦点。近期,国内各大厂商纷纷加快开展生成式AI核心技术的研发,无论是讯飞星火认知大模型展现出的雄厚“中文功底”,还是国家超级计算天津中心发布的基于国产天河超算算力、智能算力和汇集构建中文大数据集研发训练的天河天元大模型,都让大众对我国自主研发的中文生成式AI大模型充满期待。

开发适合国人的中文大模型

“AI大模型是基于海量多源数据训练的预训练模型,是对原有算法模型的技术升级和产品迭代。”国家超级计算天津中心数据智能部部长康波介绍,预训练大模型在基于海量数据的自监督学习阶段完成了“通识”教育,再借助“预训练+精调”等模式,在共享参数的情况下,根据具体应用场景的特性,用少量数据进行相应微调,即可高水平完成任务。

AI大模型能够理解人类的自然语言表达,并通过庞大的网络结构实现具有针对性的内容输出。

从效果上看,生成式AI表现为“无所不知、无所不能”,其具备了逻辑推理、上下文理解、文字创作、知识提取、代码生成等非常多元化的强大能力。

不过,目前成熟的生成式AI大模型大多基于英文数据进行训练。“通过中文或英文数据训练出来的大模型,差异还是比较大的,中文的上下文理解和语义的多解性要大于英文。大模型首先要理解人类意图,因此对于国内用户来说,用中文去训练的大模型比较适用。”康波说。

此外,生成式AI正一步步向生产工具方向发展,为产业深度赋能,或将成为人工智能与实体经济深度融合的重要力量。那么作为数据驱动的AI大模型,其训练数据来源的可靠性和安全性,便成为推动科技创新的关键。因此,自主研发中文大模型成为越来越多科技巨头的首要选择。

三月以来,国内大模型领域已进入“混战”模式,各路

玩家纷纷入局,其中有不少都“相中”了研发中文大模型。

“抢抓通用人工智能的发展机遇有几个基本要素。”科大讯飞董事长刘庆峰认为,第一,必须要在自主可控的平台上;第二,必须要同时做中文和英文,不只学习中国的“智慧”,还要向世界学习;第三,在“硬碰硬”的科技对比上,不仅要学习,还要想办法赶超。

例如,阿里推出了首个中文AI模型社区,社区首批上架超300个模型,其中中文模型超过100个,覆盖了视觉、语音、自然语言处理、多模态等AI主要领域,覆盖主流任务超过60个,且均全面开源并开放使用。360公司推出的“360智脑”背后的360GPT大模型,在海量的中文文本数据上进行了预训练和微调,从而具备了强大的语言理解和生成能力。据悉,该模型目前已经达到了100亿参数规模,并且还在不断扩展中。

中文大语言模型数据集稀缺

生成式人工智能是人工智能发展到一定阶段的产物。就像ImageNet数据集推动了残差网络等计算机视觉算法的成熟,openslr等开源数据集的发布催生了长短期记忆神经网络等自然语言神经网络的发展,图形处理器的大量使用使得模型参数从百万级发展到千亿级(ChatGPT使用了上万块A100显卡开展训练)。可以看出,生成式人工智能的快速成长,离不开算力和数据的支撑。

“大模型是大数据、大算力驱动的结果,两者缺一不可。”超级计算天津中心首席科学家孟祥飞博士强调。

一方面,中文大模型的处理能力来自于数据,它需要用海量数据来学习,通过自注意力和多头注意力机制来建立知识之间的联系。这就意味着,更多、更高质量的数据供给,将会带来模型网络中知识之间关系的完善性和贯通性。当用户提问到深层次或者冷门问题时,数据质量越高,AI大模型回答出正确答案的概率就越大。

“但目前中文大语言模型的数据集非常稀缺。”孟祥飞介绍,为了解决这个问题,天津超算中心搜集整理了全网的网页数据,并从中提取处理高质量的中文数据做成数据集,同时采集纳入各种开源训练数据、中文小说数据、古文数据、百科数据、新闻数据以及专业领域的诸如医学、法律等多种数据集,训练数据集总token数达到3500亿,训练打造了中文语言大模型——天河天元大模型。

另一方面,算力的供应是大模型的基础保障。大模型发端于自然语言处理领域,以谷歌的BERT、Open AI的ChatGPT和百度文心一言等大模型为代表,参数规模逐步

提升至千亿、万亿,同时用于训练的数据量级也显著提升,带来了模型能力的提高,这也代表着算力需求的指数级上升。

“而超级计算可以说是算力中的战斗机。”孟祥飞说,为了保证大模型的训练顺利,天津超算中心充分利用了天河新一代超级计算机的双精度、单精度、半精度融合计算输出能力,构建基于自主E级算力体系架构的智能计算引擎,建设人工智能大模型训练与应用系统支撑环境,特别是在中文处理方面构建了中文大模型数据处理的工作流技术体系,从而保障了训练任务的顺利开展。

技术成果广泛应用于多领域

在此次世界智能大会上,随着讯飞星火认知大模型一起展示的还有多款搭载了大模型的行业应用成果。

康波认为,人工智能是驱动新一轮科技革命和产业变革的巨大力量,应将大模型作为产业智能化升级的基座,用专业数据集打造更贴合行业领域的智能化高水平“专家”。

以讯飞星火认知大模型为例,该大模型的整体布局为“1+N”体系。其中“1”是指通用认知智能大模型,“N”就是大模型在教育、办公、汽车、人机交互等各个领域的应用。例如在教育领域,作为全球首款搭载认知大模型的学习机,科大讯飞推出的学习机可像真人教师一样与3岁至18岁的学生进行互动式辅导;在办公领域,基于大模型能力升级的产品具备语篇规整、会议纪要、一键成稿等功能。

康波认为,在各行各业的应用中,中文大模型显然更“接地气”。他举例说,天津超算中心综合实现了文本、语音、视频等多模态的大模型生成能力,从而形成了“一平台三能力”的基础架构,实现了更广泛的产业融合能力。基于其自然语言的理解和表达能力,与医疗结合,学习医学指南等专业规范,可以迅速地掌握对应的专业知识。其中,中文大模型可以解决“同词不同义”在医疗上的歧义性,实现精准的输出,为医疗辅助诊断提供更为全面的支撑能力。

同样,在工业检测和流程控制方面,大模型基于多元化输出能力,可以进行规范辅导、缺陷检测、流程指令生成一系列操作,降低错误率,提升生产效率。其中,中文大模型可以更好地理解复杂的专业术语以及流程指令逻辑,让输出更准确、严谨。

“在大模型通用性、泛化性以及降低人工智能应用门槛的优势推动下,人工智能也将会加快落地,形成新的机遇。”康波表示。

自对齐能否驾驭“狂奔”的AI

◎本报记者 翟冬冬

自Open AI推出ChatGPT大模型以来,AI大模型进入到公众视野,探讨热度飙升,百度、腾讯、阿里等互联网巨头也纷纷公布了其对于AI大模型的投资与布局路线图。

然而,在一片火热中,也有不少科技巨头表达了对AI大模型的担忧。利用人工智能写论文、制作虚假图片……在给人类带来便利的同时,“稚嫩”的人工智能技术也暴露出许多安全隐患问题。据媒体报道,被誉为“人工智能教父”的杰弗里·辛顿已从谷歌辞职,只为完全自由地谈论AI大模型等人工智能技术所带来的危险。

狂野生长的AI大模型,真的管不住了吗?来自美国卡内基梅隆大学语言技术研究所、国际商业机器公司研究院等机构的研究人员提出了一种全新的方法——自对齐。该方法结合了原则驱动式推理和大模型的生成能力,从而使用极少的人类监督便能实现AI大模型的“自我监管”。

和人类价值观“对齐”

目前,即便是最先进的AI大模型,也无法实现自我监管,仍需要依赖人类指令以及注释进行相应的调整。

“生成式人工智能可以被看作一个预测工具,它主要依靠大量数据样本来预测正确答案。”清华大学交叉信息研究院助理教授、人工智能国际学术交流项目主任于洋表示,所有基于数据统计的人工智能模型,对于“答案”的预测都必然有一定的随机性,而这种随机性所产生的偏差一旦与人类价值观、伦理观、道德观以及法律法规等相悖,就需要对该模型进行监管、完善。计算科学工作者致力于消除这类偏差,尽量使之与人类的价值观、伦理、法律、公序良俗等相符,这被称为“对齐”。

然而,从统计学的角度看,AI大模型一定会“犯错”。因此,AI模型对齐的监管,不应追求获得一个不犯错的完美AI大模型,而是要管理AI大模型产生违规错误的风险。“例如,对话类AI大模型在其对话中,是否会大量出现针对某个地区的地域歧视。”于洋举例道,对于AI大模型的监管,就是为了避免这类内容出现的概率过高、避免模型的答案系统性发生整体倾向这类偏误。

然而,如果AI大模型严重依赖人类监督,会造成成本过高、可靠性不足、存在偏向性等问题。因此,以最少的人类监督实现AI大模型的自我监督成为构建AI大模型的关键。

为了解决这一难题,研究人员提出了一种自对齐的新方法,仅需一个人工定义

的小型原则集(规则),便能引导AI大模型生成答复的正确导向,且能显著减少对人类监督的需求,使其几乎无需任何额外注释。AI大模型可以自对齐的设计目标是开发出有用、可靠且符合人类道德观念的AI大模型,包括在生成答案时可以反对用户的有害询问,并且为表示反对的背后原因提供解释。

值得一提的是,此前一些AI大模型至少需要5万条注释,而整个自对齐过程所需的注释量少于300行,监督效率非常高,极大地降低了AI大模型对人类监管工作的需求。

把“打骂”变为“讲道理”

那么,未来AI大模型是否可以实现完全的自我监管?于洋表示,AI大模型监管的主要方式必然是自动化、机器化、算法化的,因为AI大模型依赖海量数据支撑,仅凭人工监管是不现实的。但完全的自我监督不太可能实现,这是因为即使AI学会了相应的人类规则,但它和人类所掌握的规则数量仍然无法比较。同时,人类的思维以及逻辑推理过程具有极高的复杂性,也是AI大模型难以真正“理解”人类的原因之一。

于洋举了个例子,现在的AI大模型监管就好比教育孩子,现有的人工对齐技术

方案,其实都是在“打骂”孩子,也就是用奖励和惩罚的方式训练AI什么能做、什么不能做。而真正的教育,是要对孩子讲道理,了解为什么有些事情不能做,即要让AI理解背后的逻辑,理解人类的评判标准。这是一件很难的事情,目前大多数的AI大模型的训练和对齐技术的开发,还停留在“打骂”阶段,要从“打骂”变成“讲道理”,还有很长的路要走。

“除了技术人员开发教育AI合规的技术外,监管当局也必须制度化、系统化地从社会合规和社会影响等方面对AI大模型进行多维度的审计和监管。”于洋进一步解释道,相关行业及管理部门,需要对AI大模型违背人类价值的概率,有一个系统化的评估,并根据AI大模型相关产业的发展现状,调整监管手段,有针对性地提出监管方案,建立动态、敏捷响应技术前沿变化的监管体系。企业也要建立专门针对AI大模型的监管部门,来审核AI大模型中是否包含自我监管机制,以及具体到每一个模块将带来哪些影响,以此尽量降低AI模型的违规风险。

此外,于洋还建议,人工智能的治理,也需要大量的公众参与,因为和训练模型寻找漏洞相比,在大量应用实践中去发现漏洞、完善AI大模型,显然更加高效,也能显著降低AI大模型开发成本,推动AI大模型在各行各业的广泛应用。