



AI模型Grok 3背后的华人面孔

2月23日,美国企业家马斯克宣布,其旗下xAI公司的人工智能(AI)模型Grok 3语音模式的早期测试版正式上线。在这之前,马斯克在2月17日与三名团队成员现身发布会,推出号称“地球上最聪明的人工智能”Grok 3。与马斯克并肩而坐的三个人分别是xAI公司首席工程师伊戈尔·巴布施金、吉米·巴和吴宇怀,其中两张华裔面孔备受关注。



主笔 王晓莹

“人工智能之父”欣顿的学生

xAI公司成立于2023年7月,目标是“了解宇宙本质”。该公司官网显示,xAI的团队成员曾在知名人工智能研究实验室DeepMind、美国开放人工智能研究中心(OpenAI)、谷歌研究院、微软研究院、特斯拉及加拿大多伦多大学供职,当时该公司还在旧金山湾区招募技术人员。此次坐在Grok 3发布会“C位”的吉米·巴和吴宇怀,就在xAI公司12名初创成员之列,他们都毕业于多伦多大学。“人工智能之父”、去年的诺贝尔物理学奖得主杰弗里·欣顿就在这所大学任教多年,他的学生苏茨克维尔曾是OpenAI的首席科学家。

吉米·巴的个人资料显示,他本科、硕士、博士均毕业于多伦多大学,2011年本科毕业,2014年硕士毕业,同年他与迪德里克·金马合著论文,开发了训练深度学习模型的首选算法之一“Adam优化器”。截至2020年,这篇论文被引用超过22万次,是2016年至2020年期间被引用次数最多的科学论文。2016年,吉米·巴获得“脸书”机器学习方向的博士生奖学金,2019年博士毕业。

他的老师都是业内“大佬”,包括“深基因组学”公司创始人布伦丹·弗雷、苹果公司首任AI总监鲁斯兰·萨拉赫丁诺夫以及杰弗里·欣顿,而前两人曾是欣顿的学生。吉米·巴读博时师从欣顿,研究重点是深度神经网络的高效学习算法开发,还涉足强化学习、自然语言处理和人工智能等领域。

目前,吉米·巴是多伦多大学计算机科学系的助理教授,带过不少学生,吴宇怀就曾是他的学生。

吴宇怀1995年出生在中国浙江,母亲是一名老师。他从小就对数学和思维训练



杨格



戴子航



张国栋

的题目很感兴趣,初中毕业后,他前往加拿大读高中,后来就读于加拿大纽布伦斯威克大学和多伦多大学。吴宇怀的小学班主任方艳云用“稳重”“自律”“上进”来形容这名学生,“他在班里像个‘小大人’一样,刚入学就让他当了班长。他性格温和,和同学的关系相处得也很好”。

吴宇怀曾在朋友圈回忆初中时“死磕”一道几何题,添了8条辅助线才成功破解。当时老师也没有标准答案,还邀请他去给各位数学老师讲解答案,这让他很有成就感。2015年,吴宇怀以满绩点的成绩从纽布伦斯威克大学毕业。2017年,他获得博士生奖研金。2021年,他博士毕业于多伦多大学机器学习专业,也曾师从欣顿,还在DeepMind和OpenAI实习过,后来在谷歌任职并同时在斯坦福大学做博士后研究。

xAI公司成立后,吴宇怀加入初创团队,至今他的社交媒体X的账号还置顶着xAI成立时发布的公司目标:“解决数学问题,了解宇宙本质。”

初创团队还有三张华人面孔

xAI公司成立时就有人发现,12名初创团队成员中,有5人是华人面孔。除了吉米·巴和吴宇怀,还有哈佛大学数学专业的高材生杨格、本科毕业于清华大学的前谷歌科学家戴子航、DeepMind前科学家张国栋。

杨格出生在中国湖南,初中时赴美读书,在哈佛大学期间曾两次休学,并在休学期间追求音乐梦想,但这些都未耽误他获得哈佛大学的数学和计算机双学位。杨格的老师是1982年菲尔兹奖得主、著名数学家丘成桐,后来丘成桐和另一位菲尔兹奖得主迈克尔·弗里德曼一同向微软公司推荐杨格,后者曾在微软雷蒙德研究院担任高级研究员,与苹果、谷歌、英伟达等多家知名科技公司有过合作。

同,AI领域的先发优势并不绝对。

德国《经济周刊》发文指出,美国的科技制裁是一条错误的道路。DeepSeek的程序员没有顶级芯片,但他们用创造力弥补了这一劣势。英国《卫报》网站刊文说,DeepSeek-R1可以免费使用,并且是开源的。低成本与开放性的强强联合可能有助于普及AI技术,让其他国家(尤其是美国以外)的开发者能够入局。

xAI公司成立后,杨格曾在社交媒体上发文说:“数学赋能人工智能,人工智能助力数学研究。”其个人主页显示,他正在开发一个用于理解大型神经网络的框架,最终目标是建立一个大规模深度学习的“万物理论”。

戴子航和张国栋都是在国内完成本科学习后前往国外深造的。戴子航本科毕业于清华大学计算机科学系,2020年博士毕业于卡内基梅隆大学计算机系,曾在谷歌实习,毕业后加入谷歌研究院。他提出了多种提升AI语言模型训练效率的方法,在提高AI语言理解能力方面做出了很大贡献。

张国栋2013年考入浙江大学信息工程专业,曾获浙江大学最高奖学金荣誉竺可桢奖学金,在微软亚洲研究院实习过。2017年本科毕业后他前往多伦多大学攻读硕士和博士学位,2022年毕业后加入DeepMind。他的研究方向是训练、调优和对齐大型语言模型,2022年获得由苹果机器学习研究中心设立的苹果奖学金,当时共有15名获奖者,张国栋是4名华裔获奖者之一。

马斯克宣称,Grok 3是“地球上最聪明的人工智能”,该模型引入了包括图像分析和问答在内的高级功能,支持社交媒体平台X上的各种功能。2月23日和25日,Grok 3又对其语音模式进行了更新。

Grok首个版本发布于2023年11月,得名于科幻小说《异乡异客》中的一个火星术语,指对某事有非常深刻的同理心或直觉、能充分理解的一种状态。xAI公司当时在官网上介绍说:“Grok会带着一丝机智和叛逆精神来回答问题。如果你不喜欢幽默,请不要使用它!”去年8月,xAI公司推出聊天机器人Grok 2。马斯克称,此次的Grok 3使用了拥有约20万个GPU的大型数据中心进行训练,其计算能力是上一代版本Grok 2的10倍。

国际投行摩根士丹利表示,DeepSeek展示了一条与之前大模型不同的高效训练途径。DeepSeek的模型现在低成本优势突出,与国际知名大模型相比,其成本大约低了一个数量级。高盛集团也认为,DeepSeek新模型的成本远低于现有模型,这意味着开发利用大模型的门槛降低,互联网巨头将面临初创公司的潜在竞争。

(资料来源:新华社)

(上接A16版)

日本《朝日新闻》网站援引有“日本AI研究第一人”之称的东京大学教授松尾丰的话报道说,他试用后认为DeepSeek-R1是一个性能非常高的模型,不逊于OpenAI的ChatGPT。松尾丰说,虽然DeepSeek-R1模型并未采用什么革命性的技术,但给人的印象是做得非常好。它组合了各种好的

方法,以简单的方法实现了高性能,这一点非常重要,DeepSeek的成果证明了这样的模型也能在全球范围内竞争。

德国《世界报》以《这场AI地震始料未及——DeepSeek为何能改变一切》为题报道称,DeepSeek成功打破了AI行业此前的一种模式:如果想参与AI竞赛,就必须预先投入数十亿美元建设数据中心。美国企业试图用巨额投资构筑“护城河”,但与传统行业不