

月之暗面:

打造人人可用的通用智能助手

K 热点透视 redian toushi

6月12日,北京月之暗面科技有限公司(以下简称“月之暗面”)发布并开源Kimi K2.7 Code 编程模型,参数量达1.1万亿,提供256K上下文窗口。这一模型重点提升了长上下文编程场景的指令遵循能力、长程编程任务的性能表现,并且大幅改善了在长程任务中的过度思考倾向,平均token消耗减少30%。

月之暗面公布的基准测试结果显示,K2.7 Code 在多项编程和智能体(Agent)基准测试中较K2.6实现大幅度提升,提升比例从10%到31.5%不等。

Kimi K2.7 Code 的发布,是这家创新企业在通用人工智能(AGI)赛道上的又一次技术迭代。创立3年多来,月之暗面已经发布多款通用人工智能模型及版本,推动人工智能向深度实用化演进。目前,月之暗面已成为北京数智产业极具代表性的企业,也是我国最受关注的通用人工智能企业之一。

专注基础模型优化

Kimi K2.7 Code 编程模型的基础版,是2025年上线的万亿参数基础模型Kimi K2。值得注意的是,Kimi K2模型参数总规模虽达万亿级别,但通过工程架构优化,其激活参数仅为320亿。这种设计意味着模型能够在较低的算力成本下,开展更加复杂的计算与推理工作。

低成本与高效率这对看似矛盾的目标是如何同时达成的?月之暗面的答案是“从底层技术找突破口”。

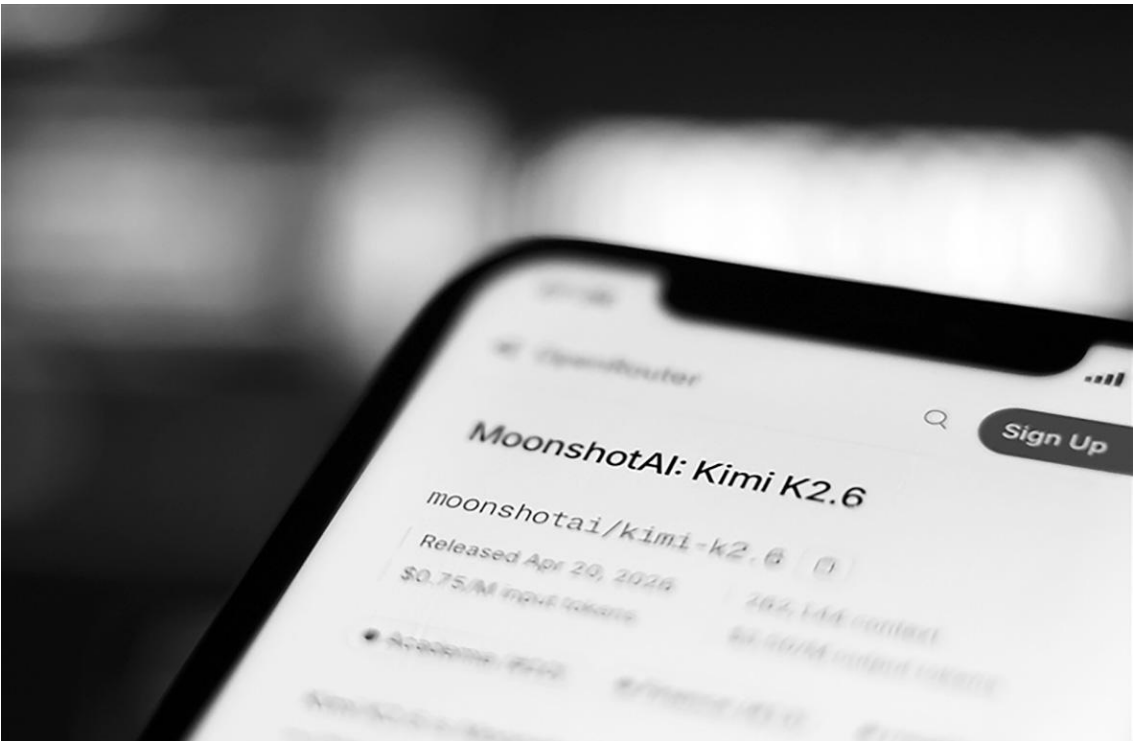
月之暗面总部位于北京市海淀区知春路。2023年4月,在生成式人工智能技术引发全球科技变革的背景下,拥有10年自然语言处理(NLP)研发经验的杨植麟与团队成员一同创办了这家企业。

“公司目前拥有约300人的科研与工程团队,其中‘90后’员工占据了一大半。”月之暗面人力资源部门负责人告诉笔者。

在这个精干的集体中,算法和基础设施领域研发人员的专业方向涵盖了自然语言处理、计算机视觉、强化学习、高性能计算等多个交叉学科。团队成员不仅参与过国际主流深度学习框架的开发,更在超大规模计算集群的运维和性能优化方面积累了丰富的工程经验。

这群年轻的科研人员普遍奉行“用时间换空间,用积累换突破”的研发哲学。在大模型行业初期热衷于拼参数规模、急于抢发过渡性案例的风潮中,团队选择保持战略定力,将主要精力倾注于基础模型的底层架构优化。

这种对底层技术创新的执着,体现在对“无损长文本”技术的攻坚上。2023年10月,Kimi智能助手首发上线,“长文本处理”成为其核心标签。



月之暗面展示 Kimi K2.6 模型

视觉中国供图

月之暗面 Kimi 研究员杜羽伦将长文本能力形象地比作计算机的“内存”。“内存越大,大模型能同时处理和理解的信息就越多,能做的事情自然就越复杂。”在当时的技术路径中,团队拒绝了众多业内常见的“技术捷径”,坚持对底层数据进行严格的无损压缩。

市场的客观数据验证了这种底层创新的价值。在公司创办一年后的2024年,Kimi的用户量便实现了百倍的增长,从最初的几十万跃升至几千万。用户群体的迅猛发展,为企业后续推出Kimi开放平台、浏览器助手以及深度研究模型提供了充足的空间。

技术谱系加速迭代

注重底层技术,既是月之暗面的初心,也是企业在发展历程中做出的动态选择。回顾Kimi核心基座模型的研发历程,2025年是企业关键的战略调整期。

2025年,面对大模型行业激烈的流量竞争以及用户增长放缓的压力,杨植麟作出决策——缩减大规模市场投流,回归通用基座模型的研发。

这一决策在研发团队内部引发了共鸣。“我们并没有急于将强化学习进行到底,而是选择回到第一性原理,先把基座模型做好。”杜羽伦介绍,团队保持定力,拒绝了抢发过渡性模型的诱惑,重点攻克超大规模智算集群的运维与架构优化。

2025年7月,这一布局结出硕果——月之暗面正式开源发布了全新基座模型Kimi K2。该模型采用混合专家(MoE)架构,在工程上实现了能效平衡。Kimi K2发布后,迅速引发了国际开源社区的关注。

此后,K2技术谱系的迭代速度明显加快。今年1月初,团队推出并开源了K2.5模型,主打代码及视觉理解;4月发布的K2.6版本,扩容Agent集群能力,该模型登顶开源模型第一;6月发布的K2.7新版本,强化长代码工程与多工具Agent能力。

杨植麟表示,团队将在做好风险控制的前提下持续突破,希望能在未来十年、二十年内,持续推出有竞争力的模型。“如进展顺利,K3系列模型可能在今年暑期就会上线。”杨植麟透露。

杨植麟曾将AI研发比作攀登一座“无限的山”。正是这种对底层技术研发的坚持,为多模态演进和长程任务执行积累了核心技术储备。

营造共享开源生态

与众多企业致力于针对特定垂直行业开发定制化软件,培养“超级员工”的思路不同,月之暗面自创立之初,其研发重心就锚定通用泛化能力。

“通用技术更复杂,有更多瓶颈,是一条更难走的路。”杨植麟坦言。

在团队的构想中,未来的大模型应当

具备跨领域的自适应学习能力,即模型在某一特定场景中学会新技能后,能够将沉淀下来的逻辑与知识自主反哺到其他核心能力上,从而走向真正的全能与通用。

与采取封闭策略的部分技术路线不同,月之暗面选择将研发成果开源。团队研发人员表示,此举旨在降低广大开发者和中小企业调用高质量AI能力的门槛,让开源生态中的每一个独立个体都能享受技术红利,共同推进应用层面的创新。

这一共享举措的背后,源自企业对通用人工智能长远发展图景的思考。

在杨植麟看来,一个模型最终如何使用不应当仅仅由单一的科技公司来定义,因为每个个体、每个企业都有其独特的应用偏好与场景需求。“每个人都应该有个性化定制的机会。”

未来,月之暗面将持续寻找将计算能源转化为高质量通用智能的最优解,致力于实现“把代码变成人人可用的工具”的愿景。同时,在产品侧,企业将更强调“与用户共创智能”的理念。

让尖端技术褪去神秘面纱,打破“千人一面”的智能化技术壁垒,让每一位普通人都能拥有专属、全能且普惠的通用智能助手。这既是这家诞生于中关村的年轻企业攻坚AI前沿的初心,也是鼓舞其在探索人工智能极限的征途中持续前行的信念。

孙明源

为工业打磨设备装上“脑眼手”

在最近举行的第十二届中国(上海)国际技术进出口交易会上,索菲丝智能科技(上海)有限公司(以下简称“索菲丝”)展出的“脑眼手”一体化智能打磨系统,吸引了与会者的目光。

换产周期最快10分钟,打磨合格率稳定在98%以上,效率提升40%……这一串数字,是索菲丝推动工业打磨从“自动化”向“自适应智能化”跃迁的生动注脚。

传统打磨依赖老师傅

打磨,是制造业中决定产品最终外观与精度的关键工序。

以往,在传统打磨车间里,工人穿着全套防护服,在弥漫的粉尘中作业,刺耳的噪声充斥整个空间。索菲丝项目经理林仕聪向笔者介绍,传统打磨产生的粉尘多为小于5微米的可吸入性颗粒,可直接穿

透呼吸道深入肺部;车间噪声普遍超过90分贝。工人长期在这样的环境中作业,易导致尘肺病和噪声聋。

此外,由于工艺复杂、非标准化程度高,打磨环节长期依赖成熟工人经验,自动化率不足10%,是自动化程度最低的环节之一。

如今,劳动力成本攀升、环保要求趋严、制造一致性需求增强,多重因素正推动打磨机器人技术加速渗透。据行业研究机构MIR睿工业统计,2025年中国工业打磨机器人市场规模约为0.9万台,同比增长12.6%,预计2026年出货量有望达到1万台。

然而,现有打磨机器人面临诸多瓶颈:产品一致性差、复杂曲面加工难度高、运维成本高、通用性不足。传统工业机器人只能按固定程序完成任务,遇到工件变形或装

夹偏差就“瞎”了,必须由经验丰富的老师傅反复示教调试。

解决打磨机器人的应用痛点,让它们能像老师傅一样“先看、再想、后做”,成了索菲丝的目标。从“按程序执行”到“自适应打磨”,索菲丝科研团队用6年时间攻克了三大关键技术难题,依托AI智能打磨技术及“匠心智脑”具身智能平台,构建起“脑眼手”一体化智能打磨系统,实现完整的“感知—决策—执行”闭环。

手艺活变数字化工艺

工业打磨中,黑色碳纤维(吸光)、高反光金属(眩光)等材质对视觉识别挑战极大。搭载了索菲丝自研的多模态3D视觉融合算法的高精度3D视觉设备就像眼睛,能在0.1秒内重建复杂曲面的高精度点云,实现微米级识别精度,并可实时识别工件位置、形状与表面状态。经3D视觉设备扫描后,智能打磨系统可直接生成三维数模,无需人工绘图。从打磨一个工件切换到打磨另一个完全不同的工件,只需10分钟。

索菲丝从2018年开始积累打磨工艺数据,从碳纤维到铝合金,从汽车外饰到风电轮毂,至今已形成千余套工艺参数。在此基础上,索菲丝自研的工业软件就像“大脑”一样,可根据工件三维模型和材质,实时生成最优打磨路径、转速和进给速度。

打磨最怕“下手重”伤及母材,也怕“力气小”磨不干净。搭载索菲丝自研的高响应力控模组的高柔顺机械臂及六维感知智能磨机则充当“手”,可精准执行打磨动作。其高响应力控模组采样频率达1000赫兹,压力波动控制在±1牛以内。更关键的是,智能磨机内置六维传感器,能在打磨过程中实时采集压力、震动、扭矩、温度等数据。“没有数

据的智能化都是空谈。我们是全球首家实现边打磨边采集工艺数据的企业,填补了打磨环节的数字化空白。”林仕聪强调。

打磨完成后,“眼”自动进行质量检测,结果反馈至“脑”,“脑”据此持续优化工艺参数,使智能打磨系统具备“观察—思考—调整”的类人能力。林仕聪介绍,这套系统打磨精度可达0.8毫米以内,合格率稳定在98%以上,综合运营成本较人工打磨节约40%,客户投资回报周期可控制在18个月以内。“该系统操作门槛极低,大部分普通工人经过一到两天培训即可上手操作。”林仕聪说。

如今在索菲丝,“脑眼手”一体化智能打磨系统安静运行,工位地面干干净净。

凭借“脑眼手”协同的差异化优势,索菲丝在汽车、航空航天、机械重工等多个领域打开了商业化局面。

林仕聪介绍,面向未来,索菲丝会探索两大方向:一是将工业软件和硬件搭载到人形机器人上;二是与西安交通大学合力打造打磨行业的“数字老师傅”。

从自研“脑眼手”一体的智能系统到填补打磨环节数字化空白,索菲丝用6年时间,走完了从“跟跑”到“并跑”再到部分“领跑”的艰难征程。工业打磨,曾是制造业自动化版图上最晦暗的“硬骨头”,而今,AI攻下了这个难题。当机械臂能够实时感知、自主决策、精准执行,打磨工序便从依赖个体经验的手艺活转变为可量化、可迭代的数字化工艺。

正如索菲丝创始人杜玲所展望的那样:“未来,大家可能会看到机器人打磨机器人。”在每一道被重新定义的工序背后,中国智造正在打磨属于自己的精度与底气。

王春

K 创新杂谈 chuangxin zatan

据报道,日前,西安交通大学首批20个微专业开启报名,这些微专业包括智能高分子、数智电气工程、智慧能源与虚拟电厂等。微专业是高校开设的“小而精”课程群,在主专业学习以外,围绕某一特定领域、研究方向或核心素养系统开设的一组课程,具有“小学分、精课程、高聚焦、跨学科、强实践”等特点。“智能传播”“智慧教育”“智慧航海与智能船舶”,微专业的设置是多元而丰富的。目前,利用周末时间或寒暑假研修微专业,成为不少高校学生的新选择。

微专业内容精炼,既高度聚焦,又强调交叉融合,为学生拓宽了知识边界,解锁了未来发展的更多空间。在教育部部署实施的高校学生就业能力提升“双千”计划等政策支持下,微专业在推动高等教育与产业需求深度融合、促进毕业生高质量充分就业方面正在发挥着切实的作用。

培养方式的创新,能让人才供给更加精准。在当下科技革命与产业变革加速推进的大背景下,社会急需大量复合型人才。微专业就成为一个很好的突破口,它可以快速立项、灵活更新课程内容、及时响应行业技术变化。微专业设置灵活,以轻量化的学习方式重组课程资源,使教学更贴近行业的真实发展趋势,帮助学生快速响应就业市场变化,提升就业竞争力。

聚焦学科前沿,突出融合与实用,才能培育拔尖创新人才。微专业的课程内容多为学科前沿或交叉内容,注重融合与实用,可为新工科、新文科、新医科、新农科提供实践与探索。微专业更加紧密关联战略性新兴产业,聚焦特定的新兴交叉领域或前沿方向,如“脑机接口与智能工程”“人工智能大模型”“可持续发展与碳中和”等,还有一些关于文化基础、自主发展、社会参与的核心素养类内容,如“国际组织与全球治理”“幸福心理教育”等。微专业构建的微型化、个性化、系统化的课程体系,可以灵活、开放、系统地培养学生的专业知识、能力和素质。

微专业更易打通学校与企业间的培养通道,把新技术、新产业引进课堂,让学生近距离感知行业发展趋势。微专业的教学环境相对开放,高校内外、课堂内外相融合。运用数字化、网络化、智能化等新技术构建知识体系,面向应用领域创新实践和交叉拓展,开阔学生学术视野。助力学生提升就业能力的同时,也为企业深度参与人才培养搭建了平台。例如,在某旅游学院开设的“民宿康养和数字化运营”微专业课堂上,有插画、茶艺、客房创意服务等实操。在有关民宿的微专业课上,民宿经营者和学生交流互动,不少学生在课堂上找到了就业方向。

未来,须紧扣国家战略与时代需求,以促进就业实效为导向,提升教学的针对性和实用性。如此,微专业才能实现良性发展。

“智能哨兵”守护西部陆海新通道

过去需要40分钟才能完成的巡检任务,如今借助AI仅需一刻钟——这是中国铁路南宁局集团有限公司(以下简称“国铁南宁局”)推动西部陆海新通道班列智能化升级的最新实践。数据显示,截至今年6月8日,西部陆海新通道班列累计运输已突破58.5万标箱。随着AI智能识别动态检车、无人机智能巡检、“低空天眼”立体防控系统三大核心应用在广西钦州港、防城港等枢纽落地,这条国际物流大通道有了“智能哨兵”,迎来效率与安全的双重提升。

作为西部陆海新通道铁海联运的关键节点,防城港站日均到发列车密度极高,车辆技术检查的速度与精度,直接决定着通道的运转效率。防城港车辆运用段5T运用车间动态检车员张蕊告诉笔者:“一列车有2600张配件图,传统作业模式下我们得逐张筛查,不仅累得眼睛发花,还总担心漏掉故障。”而AI智能识别动态检车系统依托成熟算法与高清图像识别技术,仅需2分钟即可自动完成单列货车的智能筛查,精准识别并标注故障位置。

这套智能识别动态检车系统自今年2月28日试运行以来,已累计检查6607列列车、分析2100万张配件照片,精准推送真实故障2.6万件。目前系统已覆盖防城港、钦州港、铁山港三大港区,计划今年9月正式投用。

效率的提升同样体现在钦州港站的货场。拥有7年外勤经验的货运员温锦锺表示,过去爬车顶检查不仅耗时费力,还存在高空坠落风险,集装箱搭扣、槽箱这些隐蔽部位更是难以看清。

如今,AI无人机智能巡检平台成了“空中哨兵”。国铁南宁局沿海铁路公司钦州港站货运副站长刘林介绍,这个由智能无人机场、高清摄像机和数据终端构成的平台,由AI算法自主规划无人机航线,精准捕捉列车顶部、集装箱衔接处等人工难以触及的点位,自动完成车(箱)号识别、偏载判定、杂物排查等关键作业。

“高清影像实时上传至平台,货运员通过影像就能及时溯源问题并落实整改,目前整列货车巡检仅需15分钟,作业效率较传统人工提升一倍。”刘林透露,该平台正在迭代升级,将新增车(箱)门未关检测、集装箱角件未入位识别等功能,进一步拓展AI智能巡检场景。

如果说效率是西部陆海新通道的“油门”,安全便是关键的“刹车”。在通道主干线——南防铁路上,AI同时扮演着“千里眼”与“预警员”的双重角色。

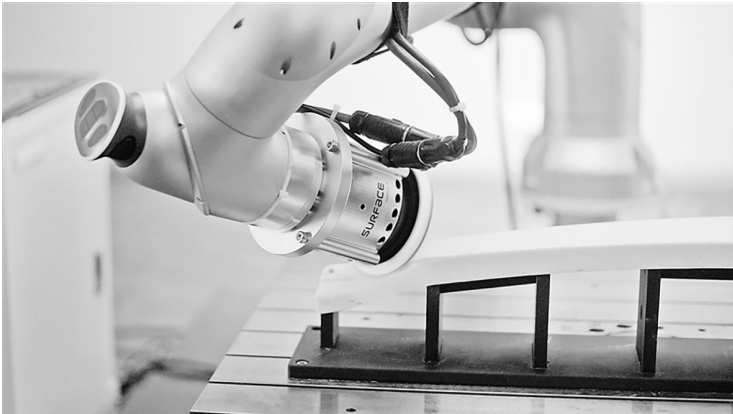
南防铁路穿越山岭河谷,地形复杂,防范压力较大。传统人工巡检长期面临“难抵达、难发现、难查全”的困境。为破解这一难题,国铁南宁局钦州工务段组建了7人专业无人机组班,配备31台高性能无人机,构建起“专业班组+定点巡查”的“天空地”立体化防控体系。

国铁南宁局钦州工务段施工办公室主任冯全良介绍,该防控体系的核心是一套名为“低空天眼”的AI无人机智能巡检系统。该系统搭载“可见光+红外热成像”双光系统,无论白天黑夜,无人机都能精准识别接触网过热、人员或机械侵入栅栏、山火等隐患,图像实时上传云端,形成“发现—预警—处置”的数据闭环。

“无人机完成巡检后,我们通过分析不同时期的图像数据,还能预判自然灾害的演变趋势。”冯全良点开一组历史数据对比图说,“这让铁路线路养护从‘事后抢修’转向了‘事前预防’的智慧运维模式。”

目前,国铁南宁局沿海铁路公司南防铁路、钦港铁路等铁路区段已投用8座无人机场,后期计划再新增34座,今年一季度该公司累计飞行1.19万架次,成功处置各类隐患395次。

韦秋莹



索菲丝智能科技(上海)有限公司的智能打磨系统通过3D视觉“看”到工件

受访单位供图