

春节档 AI 诸神之战开启

中国大模型重新定义Agent2.0时代

马年新春将至,国产 AI 大模型之战愈发火爆——短短一天时间,多家厂商密集更新模型,其中包括 MiniMax 上线新一代文本模型 MiniMax M2.5。

当各家大模型厂商围绕“春节档”展开密集发布与营销攻势,这款定位为“原生 Agent 生产级模型”的产品,正式加入这场被业界称为“AI 诸神之战”的竞争。

尽管有着编程、工具调用、办公等生产力场景全球 SOTA 成绩,这家在资本市场备受瞩目的上海大模型公司却想抛开参数规模与模型榜单的简单比拼,打开另一种叙事:中国大模型重新定义 Agent2.0 时代!

能自动处理复杂任务

在编程能力方面,M2.5 在权威榜单 SWE-Bench Verified 得分 80.2%、Multi-SWE-Bench 得分 51.3%,较上一代显著提升;在 Multi-SWE-

Bench 等多语言复杂环境中超越 Opus 4.6,达到了行业高水平。更重要的是,模型展现出“原生 Spec 能力”——在编码前主动拆解架构与功能规划,更接近真实架构师的工作模式。

工具调用和搜索能力方面,模型能够自动处理复杂任务,在 BrowseComp、Wide Search 等多项 Agent 任务中以更低的轮次消耗取得了更优的效果,相较于上一代模型表现提升 20%,达到了行业顶尖水平。

办公场景中,在 Word、PPT、Excel 金融建模等工作区间高阶场景中取得了显著的能力提升,在测评框架 GDPval-MM 与主流模型的对比中取得了 59.0% 的平均胜率。

记者获悉,MiniMax 内部已率先受益于 M2.5 的模型能力——在内部真实业务场景中,整体任务的 30% 由 M2.5 自主完成,覆盖研发、产品、销售、HR、财务等职能,且渗透率仍在持续上升。其中,在编程场景表现尤为突

出,M2.5 生成的代码已占新提交代码的 80%。

性能之外,M2.5 的另一核心亮点在于极致的推理速度与成本控制。M2.5-lightning 版本支持 100 TPS 以上输出速度,是主流模型的 2 倍左右;输入价格约 0.3 美元/百万 Token,输出约 2.4 美元/百万 Token。按每秒输出 100 Token 计算,连续运行一小时成本约 1 美元;若按 50 Token 计算,成本约 0.3 美元。这意味着 1 万美元理论上可支持 4 个 Agent 连续工作一年。MiniMax 方面认为,当性能与成本不再构成约束,Agent 规模化部署的经济模型将发生根本变化。

独创 Agent RL 支撑高速进化

过去 108 天,MiniMax 从 M2、M2.1 迭代至 M2.5, SWE-Bench Verified 成绩从 69.4 提升至 80.2,进步曲线在同业中尤为陡峭。团队将这一跃迁归因于大规模 Agent 强化学习

(RL Scaling)。

其自研 Forge 框架通过解耦训练引擎与 Agent,实现对任意 Agent 脚手架和工具的泛化优化,并通过异步调度与树状合并策略实现约 40 倍训练加速。同时,在算法层面采用 CISPO 优化与过程奖励机制,缓解长上下文场景中的信用分配问题,并将“任务真实耗时”纳入奖励函数,在效果与响应速度之间取得平衡。

据透露,M2.5 已于 2 月 12 日在 MiniMax Agent 上线,并于 13 日全球开源支持本地化部署。不到一天,来自全世界的用户已经在 MiniMax Agent 上构建了 1 万多个专家,且仍在快速增长;而 MiniMax 也针对办公、金融、编程等高频场景,在 MiniMax Agent 上构建了多组深度优化、开箱即用的专家卡组。

记者还了解到,MiniMax 将在 3 月 2 日公布 2025 年全年业绩。 本报记者 郜阳

本报讯 (记者 郜阳)近日,由上海市科创教育指导委员会联合上海市教育学会、上海市民办教育协会、长三角教育发展研究院及上海市科技艺术教育中心共同主办的 2025 上海科创教育优秀案例征集活动落幕。

本次案例征集活动以“双减赋能·科创育人——‘教育、科技、人才’一体化发展与创新”为主题,面向本市中小学、校外教育单位、科创教育基地、高校科研院所、科技场馆、科技企业等相关单位及广大教育工作者广泛征集案例,重点覆盖科学家精神引领、前沿科技资源转化、科创课程开发实践、科创教育评价探索、家校社协同育人等八大核心领域,力求精准捕捉上海科创教育的鲜活实践与创新成果。

“10+70”项优秀案例脱颖而出 勾勒科创育人的“上海样本”

评审过程中,主办方严格制定并恪守“创新性、实效性、示范性、真实性”四大核心评审标准,最终从 325 个申报案例中,评选出“十佳”优秀案例 10 个、优秀案例 70 个。

从本次评审结果不难看出,获奖案例整体呈现多元化、创新性、实效性突出的良好发展态势,覆盖多个特色领域。“十佳”优秀案例立足自身资源禀赋与发展优势,在科创教育育人模式创新、课程体系构建、实践路径探索

等方面实现重大突破,形成了可复制、可推广的先进经验,其中,上海科技馆申报的“‘发现上海’:跨学科系列研学课程实践”,紧密衔接中小学课程方案与课程标准,整合前沿科技成果,构建起科创教育基地研学活动课程体系,探索出“科技馆/基地+中小学校+高校科研院所”的跨界协同科创课程研发模式;上海市科技艺术教育中心申报的“上海市青少年科学研究院:未来科技创新英才‘三级跃迁’孵化培养体系”,精准破解未来科创英才早期培养瓶颈,为拔尖创新人才早期可持续培育提供了可借鉴、可落地的实践范式;上海农林职业技术学院申报的“‘融创联动·科普铸魂’:科学家精神立体化育人模式的探索”,创新构建“三个一”育人模式,联动多家院校及科技馆开展系列活动,破解科学家精神教育碎片化和认知表面化等难题,打造了可复制推广的科学家精神教育“上农样本”……

主办方将系统梳理并推广“十佳”及优秀案例中的先进做法,提炼可复制、可推广的科创教育“上海经验”,持续激发青少年好奇心、想象力与探求欲,为上海建设具有全球影响力的科技创新中心夯实人才基础。

中国建设银行

China Construction Bank

上海市分行

建行·沪资宝

建设资金要管好

优选建行沪资宝

70余年造价根基 铸就数字化资金监管标杆

兴建慧管

善建乐管

兴建慧管

善建乐管

兴建慧管

善建乐管