

# “AI中国团”全球霸榜，四款模型跻身前五



M2.5模型,上线不足一周便迅速登顶周调用量榜首。2月9日至15日当周,OpenRouter平台总调用量激增的3.21万亿Token中,仅M2.5模型就贡献1.44万亿Token的惊人增量。

月之暗面于1月27日发布的Kimi K2.5模型,凭借其原生的多模态架构和强大的Agent(智能体)并行处理能力,调用量实现连续跳涨。该模型能调度多达100个“Agent分身”并行工作,将复杂任务处理效率提升3倍到10倍。据媒体报道,月之暗面在发布Kimi K2.5模型后不到一个月的累计收入,已超过其2025年全年的总收入,增长主要由全球付费用户及API调用量大涨共同推动。

智谱的旗舰模型GLM-5自2月12日发布后,凭借其200K的超长上下文窗口和对长程Agent任务的深度优化,用户规模迎来高速增长,其调用量在上线次周便增长至0.8万亿Token。过去一年,阿里千问虽单个模型上榜频次不高,但知名风险投资机构Andreessen Horowitz(a16z)与OpenRouter联合发布的报告显示,其全系列模型总Token调用量以5.59万亿位居全球第二,仅次于DeepSeek(14.37万亿)。

咨询公司弗若斯特沙利文发布的报告显示,在中国大模型B端市场,2025年下半年,千问(Qwen)系列模型的日均Token调用量占比为32.1%,位列第一,相较2025年上半年的17.7%几乎翻倍,相比字节豆包(21.3%)、DeepSeek(18.4%)领先优势扩大。

对于中国AI大模型的格局,上海财经大学特聘教授胡延平在接受记者采访时提出了“AI中国团”的说法。

他认为,产业市场集中度并非越高越好,有多家头部企业形成宽广的技术产业集群,而不是少数两三家寡头,对于竞争创新和人才生态建设是好事,也有利于在中美AI竞争中形成集群优势。

a16z的合伙人Martin Casado观察到,如今在硅谷寻求融资的AI初创公司中,其路演核心模型高达80%使用中国的开源模型。

## 竞争优势: 中国Token成本不到美国AI的1/10

中国模型之所以能在短时间内席卷全球开发者,除了性能上比肩甚至超越国际顶尖模型外,其

极具竞争力的成本是另一个无可争议的核心优势。

以OpenRouter平台公示的价格为例,中国模型的成本优势一目了然。

在模型处理输入信息环节,MiniMax的M2.5与智谱的GLM-5,其价格均为0.3美元/百万Token。作为对比,海外主流的对标产品Claude Opus4.6的价格则高达5美元/百万Token,是中国这两款模型的约16.7倍。

在模型生成内容环节,成本差异更为悬殊。MiniMax M2.5的输出价格为1.1美元/百万Token,智谱GLM-5为2.55美元/百万Token,而Claude Opus4.6的价格则飙升至25美元/百万Token,分别是前两者的约22.7倍和9.8倍。

如此巨大的成本差距,直接决定了开发者在选择API时的经济考量。

这种显著的成本差异,首先源于算法层面的架构创新。

弗若斯特沙利文中国总监李庆在接受记者采访时分析指出,以“混合专家”架构为代表的技术路线,是中国模型能够大幅降低推理成本的核心原因之一。目前,包括榜单上的DeepSeek、阿里巴巴的千问3.5-Plus等模型,都已广泛采用了MoE架构。

MoE架构的巧妙之处在于,它将一个巨大的模型拆分为多个相对较小的“专家网络”和一个“门控网络”。尽管模型的总参数量可能非常庞大(如拥有数千亿参数),从而保证其“知识储备”和能力上限,但在实际处理一个任务时,门控网络会智能地判断该任务的性质,并只激活(调用)其中一小部分最相关的专家网络参与计算。

这种“按需激活”而非“全体动员”的模式,相较于传统的稠密模型(每次计算都调用全部参数),极大地减少了计算量和对硬件资源的需求。数据显示,采用MoE架构可以直接让推理时的显存占用降低60%,推理吞吐量(单位时间内处理的Token数量)提升高达19倍。这种从技术源头上实现的降本增效,是其成本优势的根本来源。

除了算法架构革新,中国AI厂商还在积极探索“垂直整合”路径,以进一步压缩每一个Token背后的成本。这条路径的核心思想,是将上层的模型算法、中层的云计算基础设施和底层的AI芯片进行深度、一体化协同设计与优化,从而解决软硬件之间

的适配痛点,榨干每一分算力。

李庆以阿里巴巴的“通义-云-芯”体系为例进行说明,这种从上到下的垂直整合模式,能够通过极致的算力调度算法,实现对底层硬件资源的最高效利用,从而大幅降低AI服务背后的基础设施成本。这种系统级优化,使得Token的生成成本得以进一步降低。

摩根大通在其研报中对中国市场作出极为乐观的预测,预计从2025年到2030年,中国Token消耗量的年复合增长率将为惊人的330%,在短短5年间实现370倍增长。

## 价值质变: Token正成为AI时代的“燃料”

Token消耗量的指数级攀升,表面看是用户规模与使用时长的增长,但其背后更深层次的驱动力,是用户对AI使用模式的根本性转变。AI的角色正在从一个提供简单信息、进行日常闲聊的“问答工具”,进化为能够深度参与工作流、处理复杂任务的“生产力工具”。

国联民生证券在近期发布的研报中,提出了“Token通胀”这一概念。这并非指Token本身变贵,而是指在单位时间内,单位用户的Token消耗结构性上升。研报将这一现象归因于三大核心趋势。

首先,用户的核心需求正在从浅层的“问答”转向深度的“干活”,即越来越多地利用AI来重构代码、改写文件、生成文档和跑测试。编程场景天然具有“长上下文、多轮迭代、大量输出”特征,会大量消耗Token。

其次,AI Agent技术的兴起和普及,放大了Token的消耗。Agent会主动规划、检索、执行、反思,多次调用模型,Token消耗自然按步骤累加。

最后是推理强度上升。更多深度思考、更长链路推理会显著提高输出与中间过程的Token消耗。对开发者而言,这往往带来更高成功率与更少返工,用户反而愿意“增加Token投入来换取效率”。

这一系列转变,意味着Token不是传统互联网时代边际成本几乎为零的“流量”,而是执行生产任务时必不可少的“燃料”。

这一趋势与全球顶尖芯片制造商的判断不谋而合。英伟达CEO(首席执行官)黄仁勋在2月26日的业绩电话会上,反复向市场强调一个核心观点:“计算即收入”“推理即收入”。他指出,没有算力,就无法生成Token;没有Token,就无法带来收入增长。在AI时代,推理性能直接决定了客户的收入能力,而推理的核心,正是高效地生成可被商业化的Token。在全球数据中心电力瓶颈日益凸显的今天,“性能/瓦特”已成为衡量AI服务效率与收入能力的关键指标。

李庆向记者表示,AI服务的商业模式正从过去单纯的“按量计费”,向“燃料+成果”的混合模式演进。一方面,作为“燃料”的Token,其单价会随着技术进步和规模效应持续下降;另一方面,随着AI从“问答”工具向“干活”的生产力工具转变,企业将更愿意为直接的“成果”付费,这将催生更多基于订阅制的商业模式。

李庆还预测,未来AI服务的定价将不可避免地走向高度定制化和灵活化。她表示,Agent时代到来意味着任务的复杂度千差万别,单一的定价模式将无法覆盖所有商业需求。未来,计算消耗、调用频次、任务是否涉及多步推理或规划等高成本操作,都将成为影响定价的因素,一个多维度、动态的定价体系将成为主流。

供稿:《每日经济新闻》作者:宋欣悦

## 榜单洗牌: 中国Token调用量首超美国

OpenRouter平台汇聚了全球数百种大语言模型,拥有超过500万开发者用户,是目前全球最大的AI模型API聚合平台。因此,其API调用量数据被视为洞察全球AI应用落地趋势最真实的“晴雨表”,因为它直接反映了开发者“用脚投票”的选择,体现了模型在实际应用中的受欢迎程度和竞争力。

值得注意的是,该平台的用户主要由海外开发者构成,其中美国用户占比高达47.17%,中国用户仅占6.01%,这使得其榜单数据更能客观反映中国AI模型在全球范围内的真实吸引力。

记者梳理OpenRouter数据发现,全球大模型Token调用量在过去一年经历了惊人的爆发式增长。2025年3月3日至9日当周,该平台前十大模型的周调用量仅为1.24万亿Token。而到2026年2月中旬,这一数字已飙升至13.95万亿Token,短短不到一年时间增长超过10倍。

2025年,美国模型是市场增长主要动力,其Token周调用量一度占据平台前十大模型总量的近七成,而同期的中国模型占比则不到两成。然而,进入2026年,美国模型的增速开始显露疲态,而中国模型则开启了“狂飙”模式。

数据显示,2026年2月第一周(2日至8日),中国模型的周调用量已跃升至2.27万亿Token,发出强烈的追击信号。

仅仅一周后,2月9日至15日当周,中国模型便以4.12万亿Token的惊人调用量,正式超越同期美国模型的2.94万亿Token,实现历史性赶超。

这股势头并未就此停止,到2月16日至22日当周,中国模型的周调用量更是冲高至5.16万亿Token,三周时间调用量增长127%,将领先优势进一步扩大。

这股强大的增长动能并非依赖单一爆款产品,而是中国AI厂商集群式崛起。

2026年2月16日至22日的周榜单显示,平台调用量排名前五的模型中,有四款来自中国厂商,分别为MiniMax的M2.5、月之暗面的Kimi K2.5、智谱的GLM-5以及DeepSeek的V3.2。这四款模型合计贡献Top5总调用量的85.7%。

具体来看,MiniMax于2026年2月13日发布的

# 汽车价格新规出台 购车透明化时代将至

争。这一转型要求企业摒弃“内卷式”竞争,通过技术突破、服务升级和品牌建设,构建可持续的竞争优势。

《指南》明确了从整车到零部件生产、从定价策略到销售行为各环节的价格合规要求,实行全流程价格管理,强化公平定价约束,规范促销与定价行为,严格规范不正当价格行为。

同时,《指南》提出,汽车生产企业除了依法降价处理积压商品外,以排挤竞争对手或者独占市场为目的实施的价格行为,存在重大法律风险。例如,整车及零部件的出厂价格低于其生产成本;通过用高规格、高等级充当低规格、低等级等手段变相降低价格,使实际出厂价格低于其生产成本;通过采取折扣、补贴等价格优惠手段,使实际出厂价格低于其生产成本等。

对于汽车销售企业,《指南》在明码标价方面作出细化规定:标价签须标明车型、配置、标准参数及可选零部件信息;销售价格为经销商自主定价,而非厂商指导价,避免指导价虚高误导消费者;不得在标价之外加价或收取任何未予标明的费用;无法现场交付车辆的,交易前须明确告知交付时间。

2025年以来,监管部门密集释放引导车市回归理性发展的信号,加码治理汽车市场非理性竞争乱象。市场监管总局在《指南》出台背景中提到,此举是为了落实综合整治“内卷式”竞争决策部署,进一步规范汽车行业价格行为,促进汽车市场健康有序发展。

中国汽车流通协会专家在接受记者采访时表示,《指南》聚焦价格行为关键环节,靶向破解流通领域突出矛盾,政策导向与行业诉求高度契合,彰显鲜明的现实针对性与前瞻性。《指南》的出台标志着汽车行业监管进入规范化、精细化新阶段,是推动产业从规模扩张向高质量发展转型的关键举措。

## 塑造市场新规则

“这几年,我每次走进4S店时,心情都十分纠结。一方面,市场上各种促销优惠的标语让人十分心动,一方面又担心‘今天买了明天就降价’。”近期购车计划的北京市消费者耿瑞的这种矛盾心理折射出当前汽车消费市场消费者的普遍心态。

自主品牌与合资品牌一线销售人员普遍反映,近两年销售任务和价格政策调整频率显著增加,新车销售利润大幅下滑,售后维修与金融保险成为主要收入来源。中国汽车流通协会统计数据显示,2025年上半年,汽车经销商亏损比例上升至52.6%,74.4%的汽车经销商有不同程度的价格倒挂。严重的价格倒挂可能导致服务缩水、售后减配等问题,最终损害广大消费者的合法权益。

记者注意到,《指南》发布后,多家自主品牌发布声明称,将持续优化自身价格管理与合规体系建设,严格落实明码标价要求,坚决杜绝价格欺诈及不正当竞争行为。

中国汽车工业协会负责人对记者表示,合法、合规是企业经营必须坚守的底线。汽车生产和销售企业应建立完善的内部价格合规管理机制,积极开展价格风险识别与预警,自觉规范非理性竞争,这不仅能够有效避免企业因价格违规导致品牌受损、客户流失等问题,帮助企业树立诚信经营的良好形象,还能增强消费者对企业的信任,在维护好消费者权益的同时,推动汽车生产企业将关注重点放在品牌、产品及服务上,从而维护公平竞争环境,促进行业健康发展。

从长远来看,核心技术是车市竞争的关键。在新能源车领域,电池能量密度、快充技术、自动驾驶等核心技术将成为竞争焦点。

中国汽车流通协会表示,部分企业长期价格倒

挂,挤压研发创新空间,破坏品牌价值与市场秩序。《指南》明确禁止低于成本销售、价格串通等不正当竞争行为,划定市场红线,将引导行业竞争从拼价格的低层次“内卷”,回归拼技术、拼服务、拼价值的良性轨道,从根本上修复产业生态。

## 购车过程更透明

对于普通消费者来说,最关心的还是新规实施后购车成本是否会出现变化。中国汽车流通协会副秘书长郎学红在接受记者采访时表示,《指南》的出台,使得接下来汽车销售价格会逐渐趋于平稳,各大车企以违背市场规律低价卖车的方式或许很难再看到,新车市场售价下降幅度和频率也会同步降低。

此外,新规将使购车过程更加透明。消费者不再需要和销售员进行复杂的价格博弈,经销商也可以把更多精力放在服务上。《指南》明确,汽车销售企业销售商品或者提供服务时,需细化明码标价规则。在经营场所及网络页面,真实准确标示商品和服务的价格信息,严禁价外加价。促销行为必须规范,显著公示促销规则、期限及限制条件,如标明赠品信息。

随着新能源汽车与智能网联技术的普及,线上交易等新业态涌现,虚假低价引流、收费规则不透明等新问题开始出现。对此,《指南》细化了价格欺诈形式,明确禁止误导性标价、虚假比价、不履行价格承诺等行为。

郎学红表示,《指南》将新业态纳入监管,明确线上平台的管理责任与价格透明要求,既保护了消费者知情权与公平交易权,也为产业创新划定了合规边界,构建了适应行业变革的长效监管机制。

据《中国消费者报》作者:吴博峰

## 价值竞争成趋势

近年来,随着新能源车快速发展及新型商业模式涌现,汽车行业价格行为日趋复杂,价格标示不规范、价格欺诈、价格串通、非理性竞争等问题频发,严重扰乱了市场秩序。

截至目前,自2023年一季度开启的汽车价格战已持续近三年,几乎席卷所有主流车企,包括自主品牌、合资品牌及新能源品牌。其中,部分自主品牌新能源车出现过一年之内三次官方下调市场售价的情况。

当价格战成为车市发展主旋律,行业整体盈利水平也持续下滑。数据显示,2025年,我国汽车行业销售利润率为4.1%,相对于下游工业企业利润率5.9%的平均水平,汽车行业利润偏低。

中国汽车流通协会乘用车市场信息联席分会秘书长崔东树告诉记者,1月共有17款车型降价,同比增加9款,新车降价潮依然在延续。其中,当月整体乘用车市场降价车型均价24.8万元,平均降价3.7万元,降幅达14.9%。《指南》的出台标志着汽车行业竞争逻辑的根本性转变:从以价格战为核心的短期博弈,转向以品质、服务、技术为核心的长期价值竞