

# MINIMAX-WP (00100.HK)

## 领先的大模型开发公司，产品商业化迅速推进

**优于大市**

### ◆ 公司研究 · 公司快评

### ◆ 互联网 · 互联网 II

### ◆ 投资评级: 优于大市(维持)

证券分析师:	熊莉	021-61761067	xiongli1@guosen.com.cn	执证编码: S0980519030002
证券分析师:	张伦可	0755-81982651	zhanglunke@guosen.com.cn	执证编码: S0980521120004
证券分析师:	陈淑媛	021-60375431	chenshuyuan@guosen.com.cn	执证编码: S0980524030003
联系人:	侯睿		hourui3@guosen.com.cn	

### 事项:

**事件:** 2月13日, 公司发布 M2.5, 在 MiniMax 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。

**国信计算机观点:** 1) MiniMax 自创立之初即布局全模态大模型方向, 聚焦基础模型研发与 AI 原生应用构建。公司通过持续迭代升级, 形成了涵盖文本、语音等能力的多模态模型体系; 2) 当前大模型能力不断提升, 应用边界快速扩展, 同时供给侧技术进步推动算力成本结构性下降, 行业从技术验证期迈入爆发增长期; 3) 公司已构建起覆盖“文本理解—视觉生成—语音生成”的多模态能力矩阵, 技术边界较为完整。这种路线使其具备长期演进潜力, 而非阶段性能力领先, 通过 C 端验证能力、B 端实现放大的双轮驱动路径, 公司调用量、用户规模快速提升; 4) 投资建议: 公司低价模型切合中小企业和个人用户需求, 多模态产品商业化和出海均处于国内领先地位, 预计公司 2026/2027/2028 年营业收入分别为 2.5/6.5/12.9 亿美元, 分别同比增长 218.7%/156.4%/100.2%, 给予“优于大市”评级; 5) 风险提示: 盈利预测的风险、AI 落地不及预期的风险、技术被赶超或替代的风险等。

### 评论:

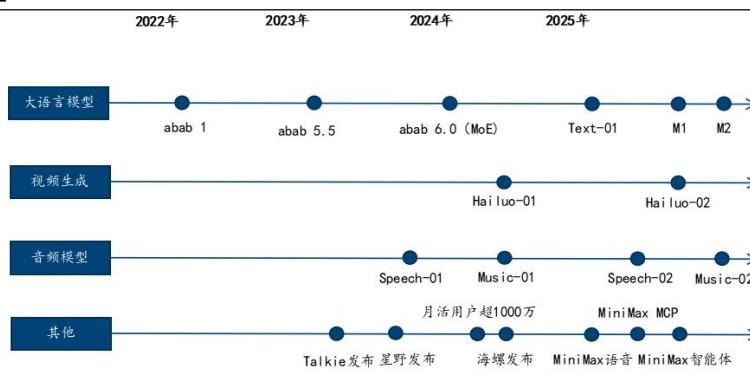
#### ◆ 领先的大模型开发企业, 兼顾模型和产品落地

MiniMax 成立于 2021 年, 自创立之初即布局全模态大模型方向, 聚焦基础模型研发与 AI 原生应用构建。公司通过持续迭代升级, 形成了涵盖文本、语音等能力的多模态模型体系, 并采用 MoE 等架构优化训练与推理效率。在模型能力基础上, 公司推出多款 C 端应用并向企业开放 API 服务, 逐步推进商业化与全球化布局。自成立之初, 公司主要经历以下发展阶段:

1) 创业初期, 产品逐步丰富 (2021-2023 年): 2021 年公司成立后便启动基础模型的开发, 并在 2022 年发布首个文本模型 abab1。2023 年发布首个百亿参数预训练大模型 abab5.5 并与首位 API 客户合作, 后续分别发布 AI 原生全模态交互平台 Talkie、全模态交互平台星野、语音模型 Speech-01;

2) 产品快速迭代, 进入快速发展阶段 (2024 年至今): 2024 年多语言语音生成模型 Speech-02、MiniMax MCP 发布。2025 年首个开源、大规模混合注意力推理模型 MiniMax M1 发布、智能 Agent 应用发布, 并发布 Hailuo-02、MiniMax M2、Music 2.0 等, 公司形成完备产品矩阵, 2026 年公司在港交所成功上市。

图1：公司产品发展历程



资料来源：公司招股说明书，国信证券经济研究所整理

### ◆ 公司为创始团队主导，管理团队经验丰富

公司核心决策权高度集中于创始团队，通过明确的职能分工实现战略、技术与商业化的协同推进。创始人兼董事会主席、CEO/CTO 闫俊杰统筹公司整体战略方向与核心技术路线，直接参与模型研发与关键技术决策。核心技术人员均拥有多年从业经验，在技术研发、经营管理与产品商业化等方面形成互补合力，共同推动公司全模态大模型迭代与 AI 产品规模化落地。

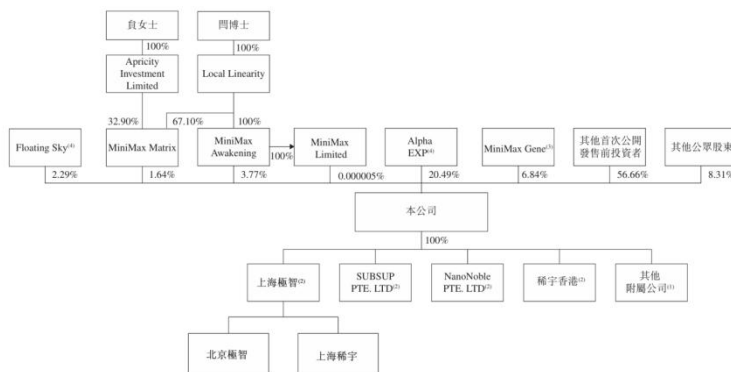
表1：公司管理层简介

姓名	年龄	职位	简介
闫俊杰	37	创始人、董事会主席、执行董事、首席执行官兼首席技术官	曾在商汤集团股份有限公司任职超过六年，担任副总裁及研究院副院长等职位；为公司创始人，主要负责监督本集团的整体管理和业务运营、董事会事务、财务事务、制定战略和运营计划(尤其是 AI 研发方面)以及作出重大业务决策。曾担任商汤集团股份有限公司融资与战略投资部经理、首席执行官行政助理及战略部总监、创新业务部总监等职位；
负焯祎	32	执行董事兼首席运营官	现任公司首席运营官，主要负责监督本集团的整体管理和业务运营、董事会事务，制定战略和运营计划(尤其是产品及商业化方面)以及作出重大业务决策。曾在北京葫芦科技有限公司(一家主要从事流媒体技术研发的公司)担任研究级软件开发工程师；
赵鹏宇	30	执行董事兼大语言模型研究与工程负责人	2023 年加入公司，现任公司执行董事兼大语言模型研究与工程负责人，主要负责与大语言模型及算法相关的研发工作。
周威聪	33	执行董事兼视觉模型研究与工程负责人	曾任职于商汤集团股份有限公司和華為技术有限公司，专注于算法领域工作；现任公司执行董事兼视觉模型研究与工程负责人，主要负责视觉模型的研发。

资料来源：招股说明书，国信证券经济研究所整理

闫俊杰及联合创始人团队通过 Alpha Exp 平台合计持有 28.25% 股份，为公司实际控制人。其余股权主要由产业资本及财务投资机构持有，其中阿里巴巴持股 13.66%，为第一大外部机构股东、米哈游持股 6.40%、IDG 资本持股 2.80% 等。

图2：公司股权架构



资料来源：公司招股说明书，国信证券经济研究所整理

◆ 公司产品：涵盖模型产品、应用产品及平台服务，形成完备矩阵

1) 基础模型：M 系列

公司构建了以 M 系列为核心的通用大语言模型体系，包括 M1、M2、M2.1 及最新发布的 M2.5 等版本。M1 为公司早期通用模型，主要承担文本理解与生成任务，在长上下文处理和基础推理能力方面奠定技术基础。MiniMax M2 在此基础上采用优化后的混合专家（MoE）架构，提升复杂推理与代码生成能力，并降低推理成本，更适用于智能 Agent 与开发类应用场景。根据 Artificial Analysis 智能指数公开评测结果，M2.5 在开源模型类别中位居前列。

图3：M2.5 在 Artificial Analysis 智能水平排序中居于前列



资料来源：Artificial analysis，国信证券经济研究所整理

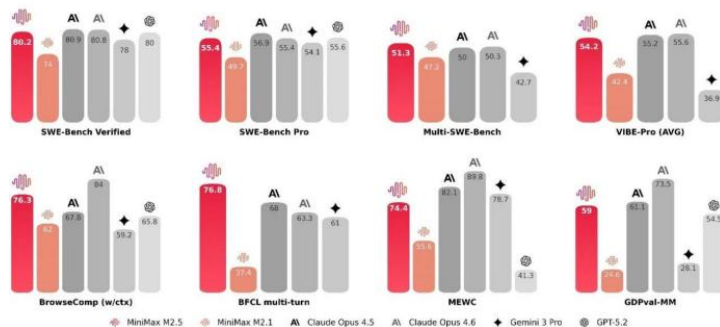
2月13日，公司发布 M2.5，在 MiniMax 内部真实业务场景中，整体任务的 30%由 M2.5 自主完成，覆盖研发、产品、销售、HR、财务等职能，且渗透率仍在持续上升。其中，在编程场景表现尤为突出，M2.5 生成的代码已占新提交代码的 80%。

- 1) 在编程、工具调用和搜索、办公等生产力场景都达到或者刷新了行业的 SOTA，比如 SWE-Bench Verified (80.2%)、Multi-SWE-Bench (51.3%)、BrowseComp (76.3%)；
- 2) M2.5 优化了模型对复杂任务的拆解能力和思考过程中 token 的消耗，使其能更快地完成复杂的 Agentic 任务。在 SWE-Bench Verified 的测试中，M2.5 比上一个版本 M2.1 完成任务的速度快了 37%；
- 3) M2.5 让无限运行复杂 Agent 在经济上可行。在每秒输出 100 token 的情况下，M2.5 连续工作一小时只

需花费 1 美金；而在每秒输出 50 个 token 的情况下，只需要 0.3 美金。按照输出价格参考，50 TPS 的版本价格是 Opus、Gemini 3 Pro 以及 GPT5 这些模型的 1/10-1/20；

4) 在编程的核心测试中，M2.5 相比于上一代模型有了显著提升，达到了跟 Claude Opus 系列类似的水平。M2.5 在超 10 种语言和数十万个真实环境中进行了训练。能够胜任各类复杂系统开发的全流程。覆盖 Web、Android、iOS、Windows、Mac 等多平台的全栈项目，包含 Server 端 API、功能逻辑、DataBase 等，而不仅仅是前端网页 demo。

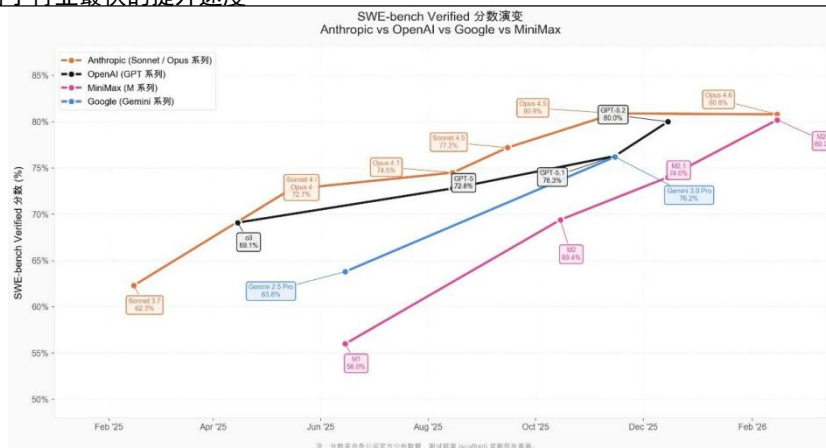
图4: M2.5 在各个测试领域取得较大突破



资料来源：公司官网，国信证券经济研究所整理

在过去 108 天里面，公司陆续更新了 M2、M2.1 和 M2.5，模型的进步速度超出公司自身预期，例如在编程领域最具代表性的 SWE-Bench Verified 上，相比 Claude、GPT 和 Gemini 等模型系列的进步速度，M2 系列模型保持了行业最快的进步速度。

图5: M2 系列模型保持了行业最快的提升速度



资料来源：公司官网，国信证券经济研究所整理

## 2) 视频生成模型：Hailuo 系列

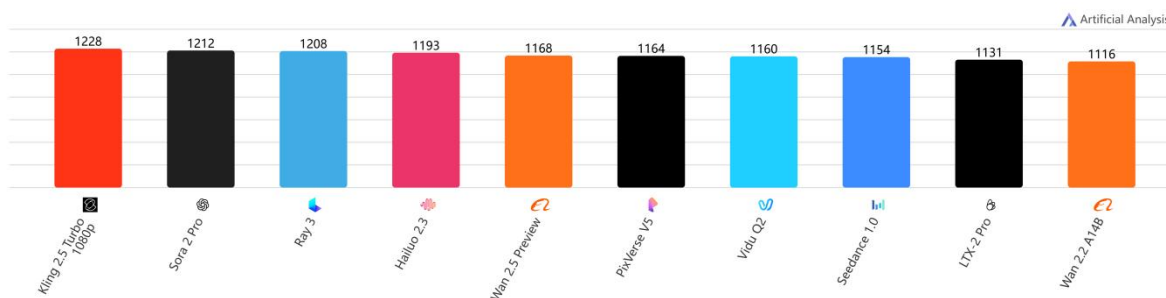
公司在多模态生成领域推出 Hailuo 系列视频生成模型，其中 Hailuo-02 支持文本生成视频及图像生成视频等多模态生成能力，在指令理解能力、复杂动作表达及画面物理一致性方面表现较为突出，能够满足内

容创作、影视辅助制作及数字媒体生产等应用需求。基于该模型，公司推出海螺 AI 视频生成平台，实现模型能力向创作端与商业化场景的延伸。

图6: Hailuo 2.3 在视频主流视频生成模型中属于第一梯队

Artificial Analysis Video Arena Quality ELO

ELO score in Artificial Analysis Video Arena (relative metric of video generation quality), Higher is better



资料来源: Artificial analysis, 国信证券经济研究所整理

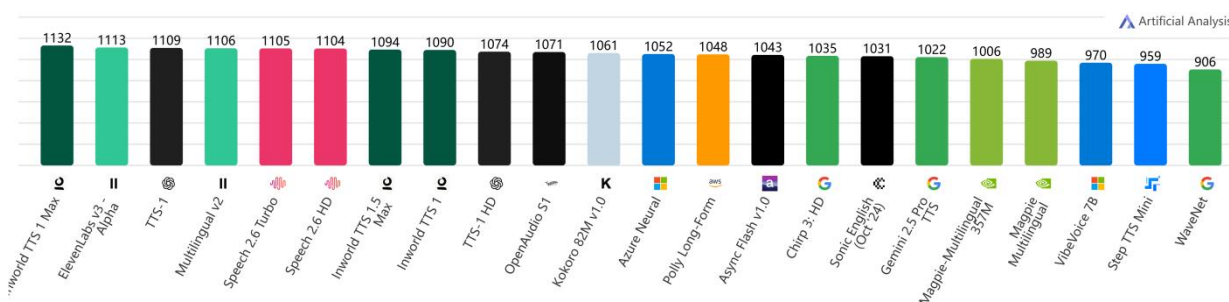
### 3) 语音模型: Speech 系列

公司于 2023 年 11 月推出首个语音模型 Speech-01, 随后于 2025 年发布 Speech-02 系列模型, 进一步强化文本到语音 (TTS) 生成能力。Speech-02 系列包括 Speech-02-HD 与 Speech-02-Turbo 两个版本, 分别面向高保真语音生成与高效率实时生成场景, 在自然度、情感表达与语音稳定性方面实现显著提升。

图7: Speech-02 在发布后处于全球领先地位

Text to Speech Arena Quality ELO

Arena ELO: Average ELO rating of the model, Higher is better



资料来源: Artificial analysis, 国信证券经济研究所整理

公司已逐步构建起模型、应用、平台协同发展的商业闭环, 以 M 系列多模态基础模型与语音、视频模型为技术底座, 通过 Talkie (星野)、海螺 AI 等应用产品以及开放平台 API 实现用户与开发者的规模化触达, 在实际使用场景中持续获取数据反馈。

1) C 端: 以订阅制与应用内付费为核心变现方式, 依托在线营销与产品分发获取用户流量, 实现高频互动

与内容消费场景变现；

2) B端：通过开放平台 API 调用收费及定制化企业服务收费，向企业客户输出文本、语音、视频等多模态能力，拓展产业级应用场景与规模化收入来源。

表2: 公司平台产品介绍

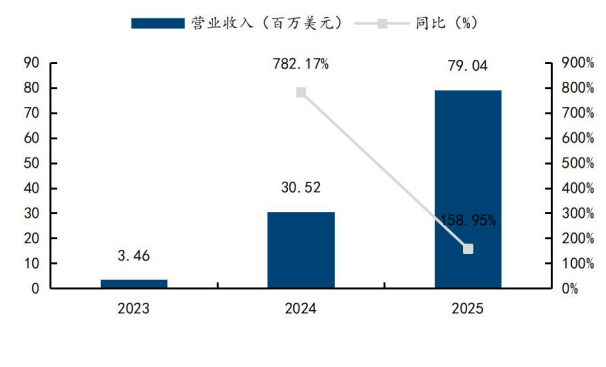
产品名称	目标用户	核心应用模型	中国发布日期	全球发布日期	主要用例	盈利模式
MiniMax	个人 (B2C)	文本、视频、语音和音乐	MiniMax Chat 发布时间: 2025 年 3 月 MiniMax Agent 发布时间: 2025 年 6 月	MiniMax Chat 发布时间: 2025 年 3 月 MiniMax Agent 发布时间: 2025 年 6 月	智能 Agent	免费增值、订阅、基于 token 的应用内购买
海螺 AI	个人 (B2C)	视频	测试: 2024 年 8 月 应用程序: 2025 年 3 月	测试: 2024 年 8 月 应用程序: 2025 年 3 月	领先的视频生成平台	免费增值、订阅、基于 token 的应用内购买
MiniMax 语音	个人 (B2C)	语音和音乐	2025 年 5 月	2025 年 3 月	音频生成工具	免费增值、订阅、基于 token 的应用内购买
Talkie/星野	个人 (B2C)	文本、视频、语音和音乐	应用程序: 2023 年 9 月 (星野)	应用程序: 2023 年 6 月 测试: 2023 年 12 月 (Talkie)	全栈虚拟交互平台	免费增值、订阅、在线广告及应用内购买
开放平台	企业 (B2B) 及开发者	文本、视频、语音和音乐	2023 年 5 月 (首个国内客户签约)	2024 年 9 月 (首个海外客户签约)	API 平台	免费增值、基于 token 的 API 计费、面向国内市场 (私有部署)、企业许可

资料来源: 招股说明书, 国信证券经济研究所整理

### ◆ 财务分析: 收入加速释放, 盈利结构边际改善

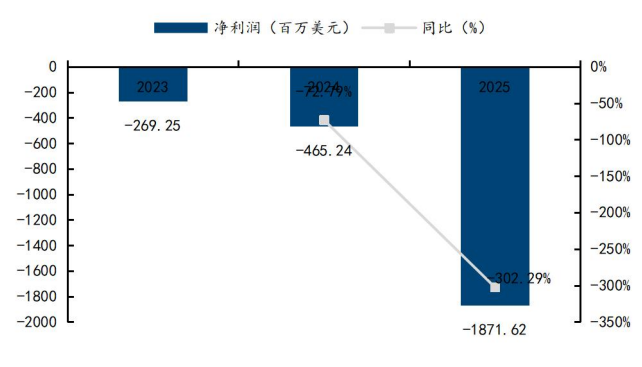
**收入规模跨越式增长, 商业化节奏明显提速。**公司自 2023 年开始形成规模化收入, 2025 年实现收入 7904 万美元, 同比增长 159%。从收入结构看, 2025 年 AI 原生产品实现收入 5308 万美元, 占比 67.2%; 开放平台及其他服务收入 2596 万美元, 占比 32.8%。2025 年实现净亏损 18.72 亿美元, 公司目前仍处于技术积累与产品商业化初期, 虽已构建 AI 原生产品矩阵, 但用户规模与商业化变现能力尚未达到盈亏平衡点。

图8: 公司收入快速提升



资料来源: 公司招股说明书, 国信证券经济研究所整理

图9: 公司利润短期承压

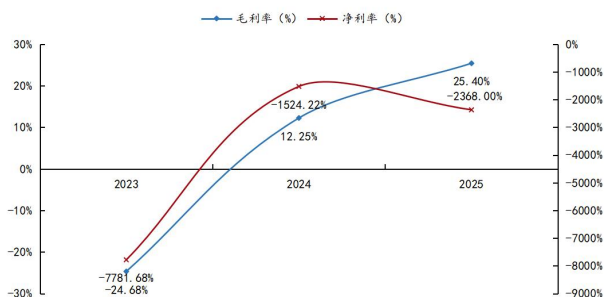


资料来源: 公司招股说明书, 国信证券经济研究所整理

**毛利率由负转正, 盈利能力稳定提升。**受模型能力提升带来的单位算力效率改善、收入结构向高附加值产品倾斜以及固定成本随收入扩大被逐步摊薄等因素影响, 公司毛利率逐步改善。2025 年公司实现毛利率 25.40%。公司净亏损逐步收窄, 期间费用率逐步下降, 2025 年公司实现销售/管理/财务费用率分别为

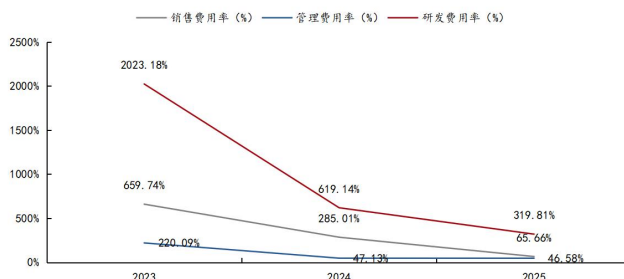
46.58/65.66/319.81%。

图10: 公司毛利率、净利率情况



资料来源: 公司招股说明书, 国信证券经济研究所整理

图11: 公司期间费用率情况



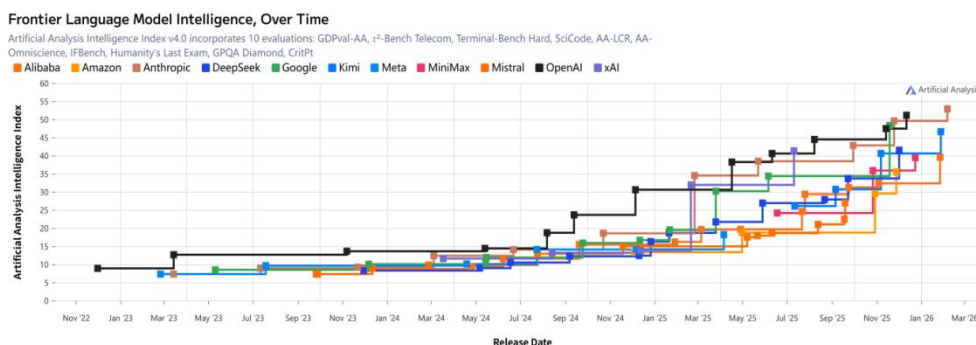
资料来源: 公司招股说明书, 国信证券经济研究所整理

### ◆ 行业趋势一: 大模型能力不断提升, 应用边界快速扩展

1) 纵向拓展: 随着参数规模的指数级攀升与混合专家 (MoE) 架构的普及, 大模型的智能水平已实现跨越式提升。以从 GPT-3 到 GPT-4 为例, 其律师资格考试成绩由倒数 10% 跃升至前 10%, 模型推理能力发生飞跃式提升。而 GPT-4o、Claude 3.7 等后续迭代更新也持续刷新了智能上限。与此同时, 模型能力的边界也在不断外延: 一方面, 长上下文技术 (如 10 万+ token) 打破了信息处理的记忆瓶颈; 另一方面, Agentic AI 的兴起标志着模型从“被动问答”向“主动执行”的范式转变——通过自主规划并调用代码解释器、浏览器等外部工具, 大模型已具备在复杂环境中解决实际问题的逻辑推理与执行能力。

2) 横向拓展: 多模态融合加速迈向 AGI, 大模型正加速突破单一文本模态的局限, 构建起涵盖视觉、听觉及视频的统一语义空间。在理解侧, GPT-4V 与 Gemini Ultra 已能精准解读医学影像与复杂图表; 在生成侧, 从 Midjourney 的艺术创作到 Sora、Veo 3 的视频生成, AI 的创作边界被无限延展。目前, 学术界与产业界正致力于构建集理解与生成于一体的统一架构, 构建起视听多模态的无缝交织的模型能力。

图12: 各前沿模型评测得分不断提高



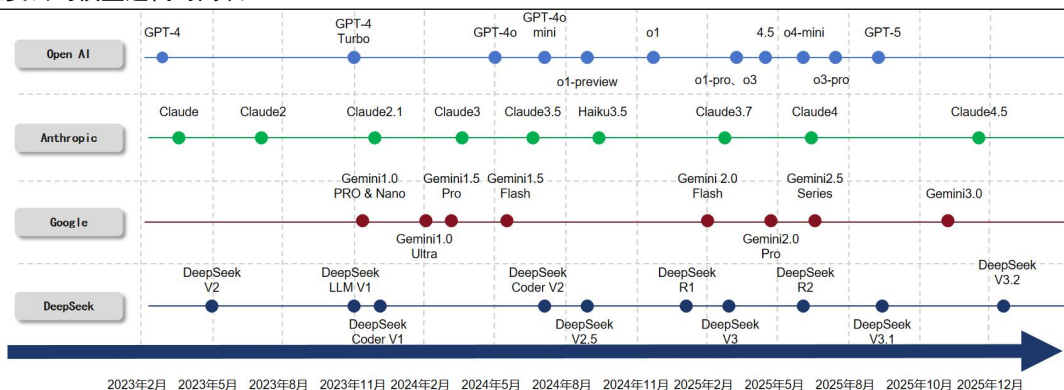
资料来源: Artificial Analysis, 国信证券经济研究所整理

模型迭代速率加快、规模与能力不断提升。2025 年以来, 全球大模型行业正式告别了以年为单位的代际演进, 转而进入了以季度甚至月度为周期的竞速时代。头部厂商的模型更新频率已从 2024 年的四个月以上大幅压缩至三个月内, 以 Anthropic 为例, 2024 年 3 月公司推出了 Claude3 系列, 随后推出 Claude3.5 系列, 并分别于 6 月和 10 月推出 Sonnet 和 Haiku 系列, 共同构成了对 Claude3 家族的重大升级。2025 年,

Anthropic 于 2 月推出 Claude3.7 Sonnet，三个月后的 5 月又推出 Claude 4，8 月公司就推出 Opus4.1，更新速度相较于 2024 年快了进 50%，从季度更新向月度更新演进。

国内方面，Deepseek、字节跳动、阿里巴巴、智谱、Minimax 等头部厂商也并未放缓其模型迭代的速度。以 Deepseek 为例，继 2024 年底发布 V3 及 2025 年 1 月推出 R1 模型后 DeepSeek 便迅速于 3 月和 5 月推出了 V3-0324 与 R1-0528 等增强版本，并穿插发布了 DeepSeek Prover-V1（4 月）以强化逻辑证明能力。进入下半年，其技术飞轮进一步加速：8 月推出 DeepSeek V3.1，9 月发布 3.2-Exp 实验版，10 月与 11 月分别补齐了 DeepSeek OCR 与 DeepSeekMath V2 等关键模态拼图，最终在 12 月正式落地 DeepSeek V3.2。模型能力亦不断提升，迭代加速。

图13: 各主要公司模型迭代时间表

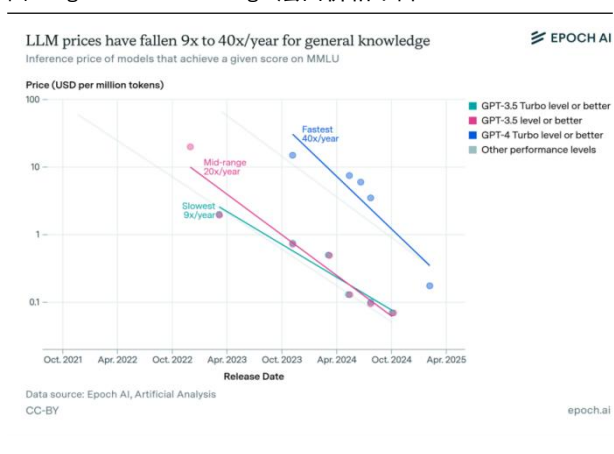


资料来源：各公司官网，国信证券经济研究所整理

◆ 行业趋势二：供给侧技术进步推动算力成本结构性下降，市场空间逐渐打开

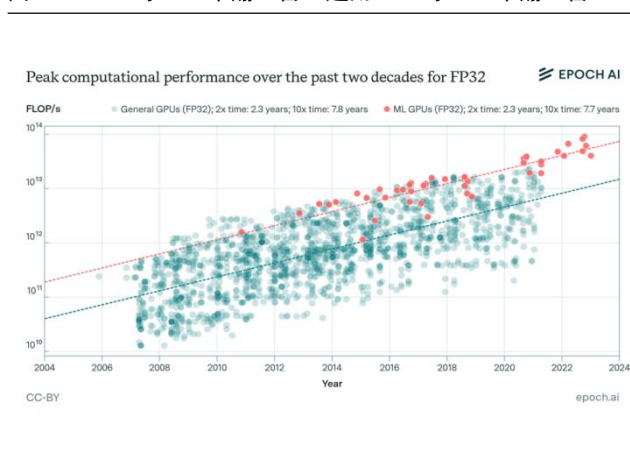
供给侧的技术进步正在驱动算力成本的结构下降，打开应用侧市场空间。据 EPOCH AI 数据，机器学习专用 GPU 的算力性价比（FLOPs/\$）平均每 2.1 年翻倍，叠加 FP16、INT8 等低精度计算与 Tensor Core 的软硬协同优化，推理效率可实现数量级提升。这种底层硬件迭代，直接导致了高性能模型推理成本的快速下探，为大规模商业化应用扫清了核心成本障碍。

图14: general knowledge 层面价格下降 9x-40x



资料来源：EPOCH AI，国信证券经济研究所整理

图15: MLGPU 每 2.1 年翻一番、通用 GPU 每 2.5 年翻一番

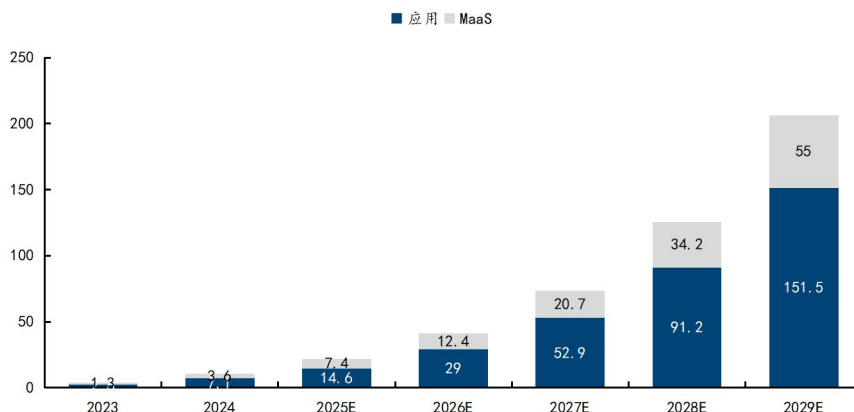


资料来源：EPOCH AI，国信证券经济研究所整理

算力成本的降低向应用层传导，表现为“每百万 Tokens”价格的持续下行。以 GPT-4 level 为例，2025 年每百万 token 成本仅为 2024 年的 1/40，使用成本显著降低。这将显著提升下游企业与个人用户的付费

意愿。随着推理成本从高昂走向普惠，高频次、长上下文的复杂应用场景（如全库代码扫描、实时伴随助手）将具备经济可行性。供给端成本的指数级下降将激发需求端的非线性增长，从而激活下游企业和个人需求，推动行业从技术验证期迈入爆发增长期。据灼识咨询数据，全球大模型市场规模，按基于模型的收入计全球大模型市场将由 2024 年的 146 亿美元增长至 2029 年的 2065 亿美元，CAGR 达 80.7%，其中大模型应用市场规模从 71 亿美元增长至 1515 亿美元，CAGR 达 84.3%。Maas 市场从 36 亿美元增长至 55 亿美元，期间 CAGR 达 72.7%。

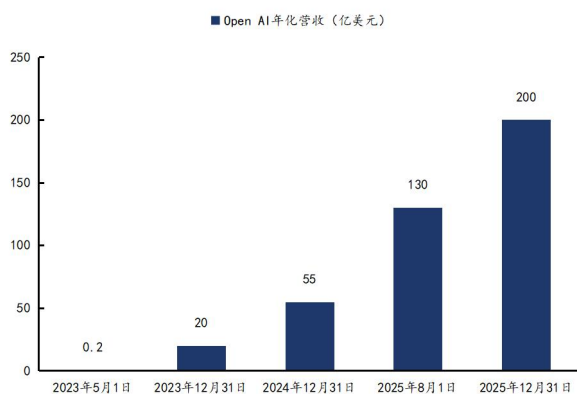
图16: 全球大模型市场规模，按基于模型的收入计（十亿美元）



资料来源：灼识咨询，国信证券经济研究所整理

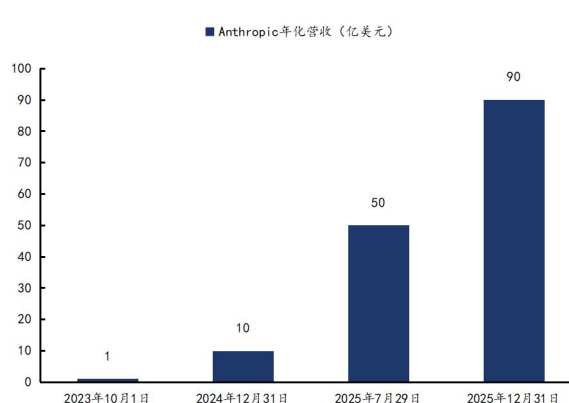
随着成本下降和应用场景拓展，头部企业已逐渐实现商业化落地。数据显示，2023 年-2025 年，OpenAI 的年化营收从 0.2 亿美元增长至 200 亿美元。同期，Anthropic 亦展现出强劲的增长爆发力，年化营收从 1 亿美元攀升至 90 亿美元。其它头部厂商如 Xai 等也同步推进，随着技术进步，成本降低以及多模态大模型的演化，未来商业化路径有望进一步拓宽。

图17: OpenAI 年化营收（亿美元）



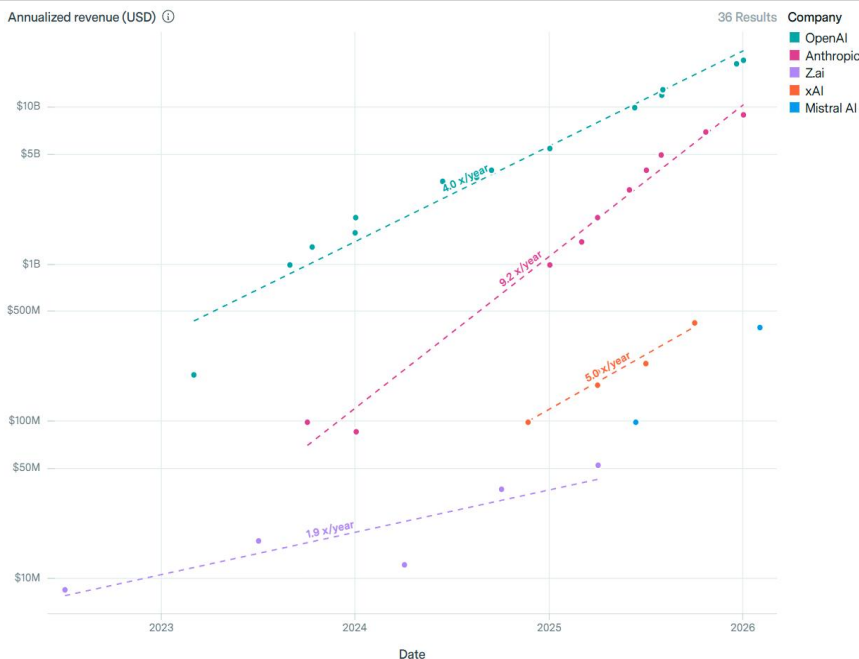
资料来源：EPOCH AI，国信证券经济研究所整理

图18: Anthropic 年化营收（亿美元）



资料来源：EPOCH AI，国信证券经济研究所整理

图19: 美国主要大模型公司年化营收（十亿美元）



资料来源：EPOCH AI，国信证券经济研究所整理

### ◆ 行业趋势三：竞争格局海外暂时领先，国内外差距逐渐缩小

目前大模型应用的主要赛道包括生产力、娱乐、视觉生成、音频生成和通用 2B 服务。纵观这些赛道，由于大模型技术的通用性特征，这些市场都实现了用一套高度可扩展的模型、撬动下游规模化与个性化并存的多元需求场景，从大型行业客户到中小创作者全覆盖，实现模型 ROI+转正。

这些赛道里能长期保持自然增长的产品，都是底层模型智能持续提升驱动的。长期头部玩家需要有可端到端优化的底层自研模型，且保持模型水平在第一梯队。一个模型智能水平的突破能使其快速突围，当技术突破带来更佳用户体验时，用户就会自然迁移。

1) 生产力：生产力赛道可以分为通用类和垂类厂商，其中通用类头部厂商包括 OpenAI、DeepSeek、Google 等，垂类头部厂商包括 Cursor 等。赛道的关键在于 1) 从 Chatbot 向 Agent 演进：**核心逻辑在于模型能力的提升**，使其不再局限于简单的问答，而是具备长程任务规划、多步推理及工具调用能力，能够处理复杂的学术研究或商业分析任务。2) 代码开发的门槛降低：随着 Claude 3.5 等模型在代码能力上的突破，应用形态从单纯的代码补全转向 Vibe Coding（自然语言编程）。这使得非技术背景用户也能通过对话构建应用，同时也极大提升了专业开发者的效率；

2) 娱乐：赛道的头部厂商包括 Character.AI、Talkie（星野）等，该赛道通过提供个性化的智能体互动，满足用户的社交与情感需求。**核心竞争力在于利用多模态工具（语音、图像）丰富交互形式，激发用户的创作欲望，从而建立高活跃度的社区生态；**

3) 视觉生成：头部玩家包括 OpenAI、Google、Minimax，其核心技术壁垒在于声音的逼真度（情感表现力）以及实时交互的响应速度。在 B 端主要用于突破人力瓶颈（如智能客服、销售 Agent），在 C 端则服务于内容创作（有声书、配音、播客生成），语音正逐渐成为大模型时代的通用交互入口；

4) 通用 ToB 服务：头部厂商包括 OpenAI、Anthropic 等，以及 Microsoft、Amazon、Google、Alibaba 等云服务商，其通过 API 或私有化部署提供算力与模型服务，**核心竞争要素回归到性能、成本（ROI）、并发稳定性及数据合规性**。企业客户倾向于根据具体场景路由不同的模型，以在成本与效果之间取得平衡，推理成本的下降是推动产业渗透的关键驱动力。

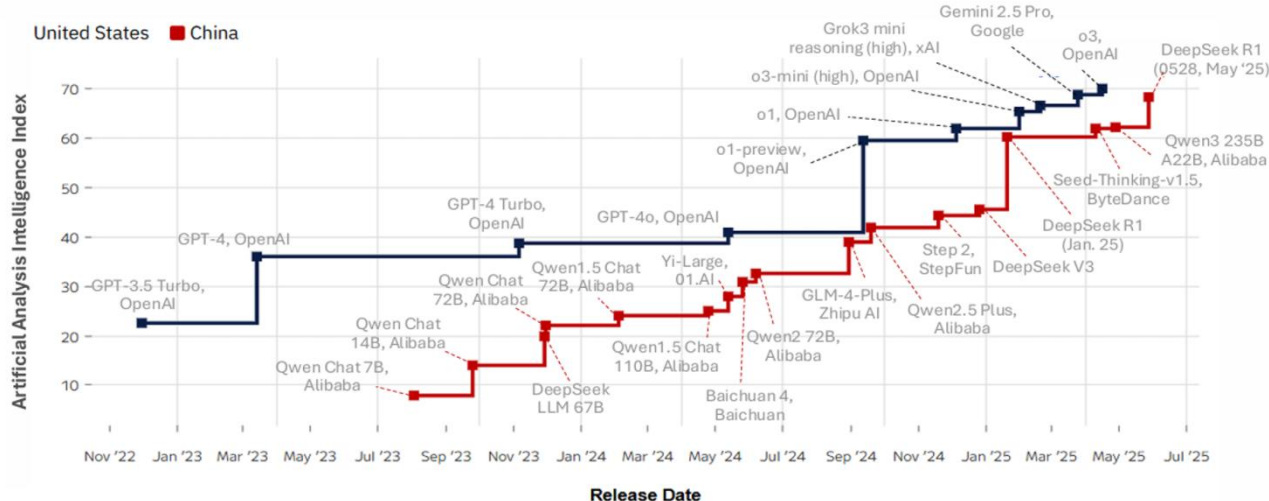
图20: 大模型应用市场格局一览



资料来源: 招股说明书, 国信证券经济研究所整理

美国顶尖公司和实验室占据模型性能高地，中美模型代差逐渐减少。自 2022 年 ChatGPT 发布以来，中美前沿模型之间的性能差距虽然一直存在，但目前已缩窄至历史最小水平。根据 Artificial Analysis 数据，中国领先 AI 实验室与美国领导者的性能差距已从 ChatGPT 发布后的超过一年，缩小到不到三个月。

图21: 中美大模型代差不断缩小



资料来源: Artificial Analysis, 国信证券经济研究所整理

◆ 公司优势: 多模态布局形成平台级技术优势, 具备长期演进能力

多模态能力布局形成平台级技术优势，公司自成立早期即明确推进多模态底层能力建设，在语言、语音、视觉及视频等方向分别构建自研模型体系。整体来看，公司已构建起覆盖“文本理解—视觉生成—语音生成”的多模态能力矩阵，技术边界较为完整。这种路线使其具备长期演进潜力，而非阶段性能力领先。在文本模态方面，公司通过长上下文建模与高效专家混合架构（MoE）优化，在复杂推理、代码生成及多任务泛化能力上持续提升，为多模态融合提供稳定的语义理解核心。在视觉模态方面，视频生成模型在物理一致性、复杂动作控制与连贯性表达上实现突破，强化了高维动态内容的生成质量。在语音模态方面，公司模型在自然度、情感表达与多语言泛化能力上达到行业领先水平，已具备规模化应用条件。

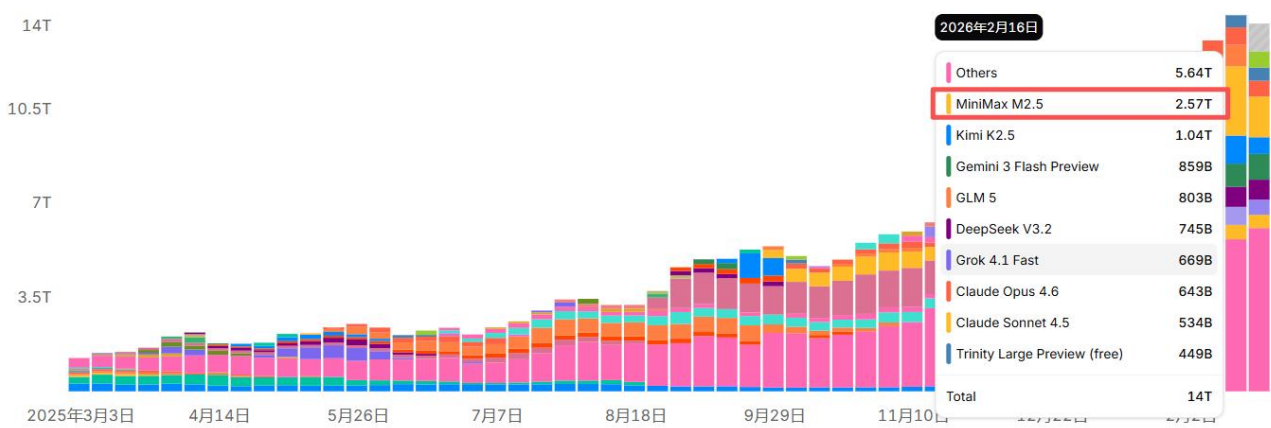
表3: 公司核心模型介绍

类别	最新核心模型	发布日期	描述	产品亮点
文本	MiniMax M1	2025年6月	基于 MoE 架构的推理模型	软件工程能力佳、擅长长上下文理解、Agent 应用能力强
	MiniMax M2	2025年10月	基于 MoE 架构的推理模型	认知和推理能力强、编程能力先进、Agent 应用能力强、效率和成本优势
	MiniMax M2.1	2025年12月	基于 MoE 架构的推理模型	大幅提升多语言编程，更能胜任复杂推理任务和办公场景的实用性优势
	MiniMax M2.5	2026年2月12日	基于 MoE 架构的推理模型	在编程、智能体任务与生产力场景中取得显著提升，进一步强化了复杂任务执行效率与成本优势
视频	Hailuo 2.3	2025年10月	视频生成	更广泛的艺术和风格化能力，精细的面部和表情建模，成本性能表现佳

资料来源：招股说明书，国信证券经济研究所整理

2026年2月，MiniMax 推出新一代旗舰模型 MiniMax M2.5，围绕真实生产环境中的复杂任务处理与多场景应用能力进行系统性升级。相较上一代版本，M2.5 在长上下文理解、复杂编程推理及工具调用协同方面进一步强化，兼顾性能提升与推理成本优化，持续拓展其在企业级生产力场景中的适配深度。据 OpenRouter 最新真实调用数据统计，MiniMax M2.5 上线后已进入全球模型调用排行榜前列，同时 M2.1 仍保持较高活跃度，双型号产品矩阵共同跻身平台前十，显示其在开发者生态与实际使用场景中的市场认可度持续提升。

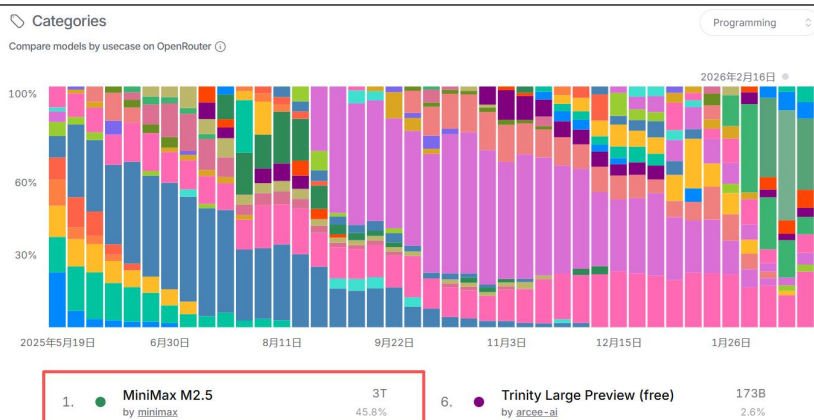
图22: M2.5 系列模型在 OpenRouter 平台上调用量达全球第一



资料来源：Openrouter，国信证券经济研究所整理

2026年2月，在 OpenRouter Programming 最新统计中，MiniMax M2.5 以约 29.7% 的调用份额位列第一，调用规模超过 1.21T tokens，显著领先同类主流模型。反映其在编程场景中的实际落地能力与生态渗透度已进入全球第一梯队，在复杂代码生成、工程级任务理解与生产力工具协同方面具备较强竞争力。同时，平台整体调用量持续增长，说明编程场景需求仍在扩张，领先模型具备更强的份额集中效应。

图23: 公司 OpenRouter Programming 模型调用份额第一



资料来源: Openrouter, 国信证券经济研究所整理

◆ 公司优势: 采取“C 端验证能力、B 端实现放大”的双轮驱动路径

公司早期以 AI 原生应用为突破口, 通过高频互动场景实现模型能力的产品化落地, 在用户规模扩张过程中完成技术与商业模式的双重验证; 在此基础上, 再将成熟的模型能力向开放平台与企业服务延伸, 形成由应用牵引、平台承接的演进路径。

**在 C 端,** 公司并未选择以通用问答为核心的竞争路线, 而是切入情感陪伴、角色互动与沉浸式对话等高黏性场景, 强调人格一致性、多轮对话稳定性与情绪表达能力, 形成差异化体验优势。该类场景天然具备高使用频率与情感依附属性, 使用户付费意愿相对工具型 AI 更强, 从而支撑较快的商业化落地节奏。同时, C 端高频真实交互为模型持续优化提供数据基础与工程验证环境, 形成技术能力与产品体验之间的正向循环;

**在 B 端,** 公司则侧重将 ToC 阶段沉淀的模型能力与工程经验进行平台化输出。角色构建、长对话一致性、多模态生成与成本控制等能力, 可以迁移至游戏、虚拟人、内容平台及企业智能化场景。前期通过 C 端打磨形成的稳定性与调优能力, 也在一定程度上降低了企业级部署与交付门槛。整体来看, B 端更强调规模化扩展能力与收入结构的稳定性, 是公司技术能力外溢与长期增长的承接方向。

图24: AI 原生产品细分收入 (百万美元)

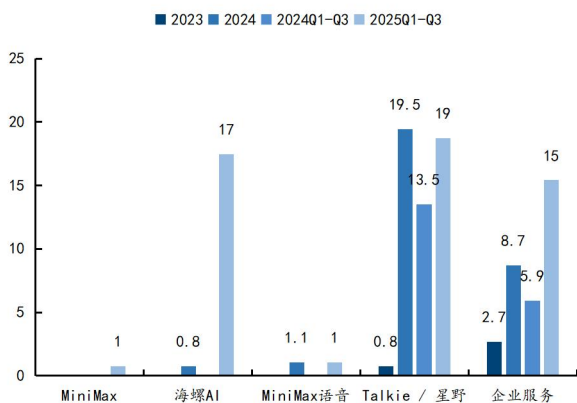
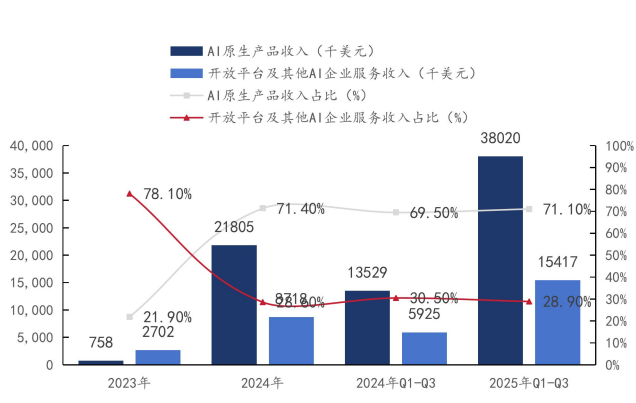
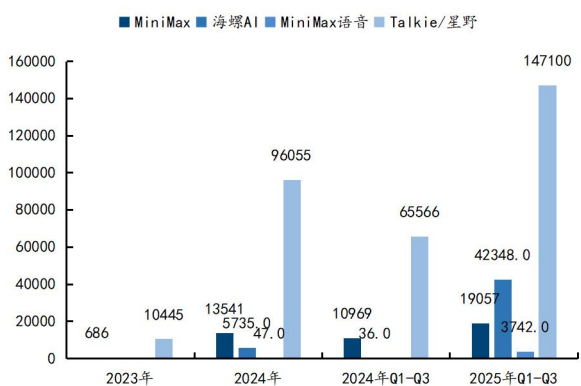


图25: MiniMax 各产品收入及占比 (千美元、%)



资料来源：公司公告，国信证券经济研究所整理

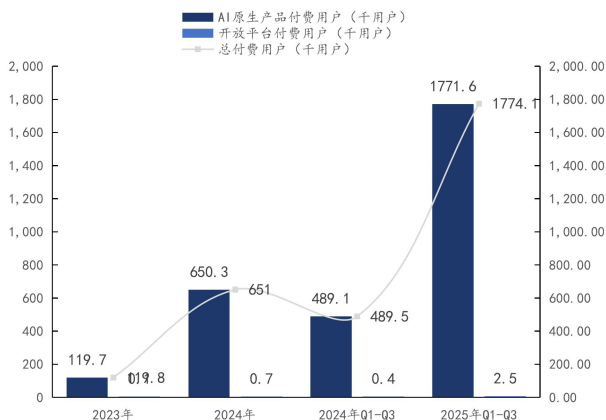
图26: AI 原生产品细分收入（百万美元）



资料来源：公司公告，国信证券经济研究所整理

资料来源：公司公告，国信证券经济研究所整理

图27: MiniMax 各产品及开放平台用户数（百万用户）

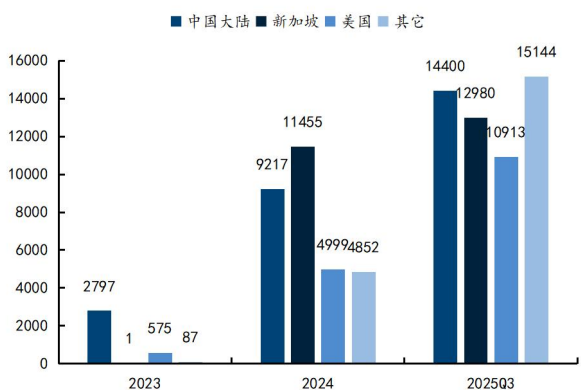


资料来源：公司公告，国信证券经济研究所整理

◆ 公司优势：全球化推进迅速，确立“首日即全球”的发展锚点

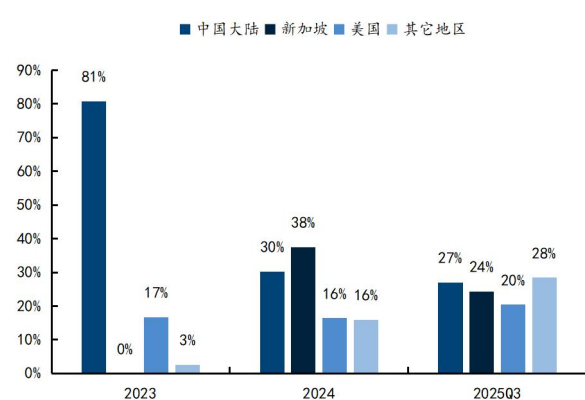
公司自创立伊始便将全球化确立为企业的核心生存法则与发展逻辑，每一款产品从研发阶段就致力于支持多语言环境，确保上线即面向全球 200 多个国家和地区发布。全球化基因使得公司能够迅速跳出单一市场的内卷，在新加坡、美国等高价值市场通过精准的本地化运营逐渐扎稳脚跟。在具体产品落地通过“情感陪伴+生产力工具”的双引擎策略，公司在全球范围内实现用户和营收双增长。旗下 Talkie 以极高的用户粘性不仅实现了日均使用时长超 70 分钟的深度互动，更构建了累计超 2.12 亿用户的全球社区；而生产力工具海螺 AI 则大幅降低了视频创作门槛，不仅支持用户创作出播放量过亿的剧集，更成为全球最受欢迎的 AI 视频平台之一。截至 2025 年前三季度，公司海外收入占比已高达 73.1%，付费用户数从 2023 年的约 12 万飙升至 177 万，且付费用户平均支出（ARPU）翻倍至 15 美元。在用户规模和平均单价方面均实现了高速增长。

图28: 公司分地区营收（千美元）



资料来源：招股说明书，国信证券经济研究所整理

图29: 公司分地区营收占比（%）



资料来源：招股说明书，国信证券经济研究所整理

### ◆ 盈利预测

我们的盈利预测基于以下假设：

**AI 原生产品（C 端）：**涵盖 MiniMax、海螺 AI、MiniMax 语音、Talkie/星野等 C 端业务，2023-2025 年占公司总收入比重分别为 22%/71%/67%，当前公司核心大模型迭代迅速，多模态业务商业化领先，公司有望继续推出更低价格，更高质量的 C 端 AI 产品，打开收入增长空间。我们预计 2026/2027/2028 财年 AI 原生产品收入增速分别为 130.0%/110.0%/70.0%，占公司总收入比重分别为 48%/40%/34%。

**开放平台及其它服务（B 端）：**涵盖 B 端基于 AI 的企业服务收入，2023-2025 年占公司总收入比重分别为 78%/29%/33%，当前全球 AI 创业需求旺盛，token 消耗量迅速上涨，公司提供全模态和高质价比的产品，满足复杂的 AI 创业需求。我们预计 2026/2027/2028 财年开放平台及其它服务收入增速分别为 400.0%/200.0%/120.0%，占公司总收入比重分别为 52%/60%/66%。

表4: 公司营业收入预测

	2023	2024	2025	2026E	2027E	2028E
<b>总收入（百万美元）</b>	3	31	79	252	646	1293
增速（%）		782.2%	158.9%	218.7%	156.4%	100.2%
毛利率（%）	-24.7%	12.2%	25.4%	32.4%	38.2%	43.8%
<b>一、AI 原生产品（C 端）</b>						
收入（百万美元）	1	22	53	122	256	436
增速（%）		2776.6%	143.4%	130.0%	110.0%	70.0%
占比（%）	22%	71%	67%	48%	40%	34%
<b>二、开放平台及其他服务（B 端）</b>						
收入（百万美元）	3	9	26	130	389	857
增速（%）		222.6%	197.8%	400.0%	200.0%	120.0%
占比（%）	78%	29%	33%	52%	60%	66%

资料来源：公司公告，国信证券经济研究所整理

**主要费率：**公司正处于高速发展期，随着公司收入规模扩大，公司研发、销售和行政费用将得到明显改善，未来研发、销售和行政费用率均随收入规模增长逐步下降。我们预计 2026-2028 财年销售费用率为 30.9%/16.9%/11.0%，管理费用率为 21.9%/11.1%/6.7%，研发费用率为 141.2%/74.5%/47.4%。

表5: 公司期间费用预测

年份: 财年	2023	2024	2025	2026E	2027E	2028E
销售成本（百万美元）	4	27	59	170	399	727
增速（%）		520.9%	120.1%	188.9%	134.4%	81.9%
毛利率（%）	-24.7%	12.2%	25.4%	32.4%	38.2%	43.8%
销售费用（百万美元）	23	87	52	78	109	142
增速（%）	3788.8%	281.1%	-40.3%	50.0%	40.0%	30.0%
销售费用率（%）	659.7%	285.0%	65.7%	30.9%	16.9%	11.0%
行政费用（百万美元）	8	14	37	55	72	86
增速（%）	137.0%	88.9%	155.9%	50.0%	30.0%	20.0%
行政费用率（%）	220.1%	47.1%	46.6%	21.9%	11.1%	6.7%
研发费用（百万美元）	70	189	253	356	481	612
增速（%）	562.9%	170.0%	33.8%	40.7%	35.2%	27.3%
研发费用率（%）	2023.2%	619.1%	319.8%	141.2%	74.5%	47.4%

资料来源：公司公告，国信证券经济研究所整理

按上述假设条件与假设，我们得到公司 2026/2027/2028 年营业收入分别为 2.5/6.5/12.9 亿美元，分别同比增长 218.7%/156.4%/100.2%。

#### ◆ 投资建议

**领先的全模态大模型提供商，在消费级和国际化领域有较大空间，给予“优于大市”评级。**1) 文本模型：M2.5 模型在编程、工具调用和搜索、办公等生产力场景均达到或刷新了行业 SOTA，在每秒输出 100token 的情况下，连续工作一小时仅需花费 1 美金，市场潜力巨大；2) 多模态：公司全模态技术领先，旗下 AI 视频、AI 陪伴、AI 音频等产线齐全，并且向企业和个人提供全模态 AI 能力，目前已有超过 13 万家企业客户和开发者接入，商业化有望迅速落地；3) 海外业务：海外收入占比 73%，且欧洲和中东地区收入迅速增长。考虑到公司低价模型切合中小企业和个人用户需求，多模态产品商业化和出海均处于国内领先地位，未来业绩增长空间有望进一步打开，我们给予“优于大市”评级。

#### ◆ 风险提示

**盈利预测的风险：**1) 我们假设公司未来 3 年收入增长 218.7%/156.4%/100.2%，可能存在对公司产品销量及价格预计偏乐观、进而高估未来 3 年业绩的风险。2) 我们预计公司未来 3 年毛利分别为 32.4%/38.2%/43.8%，可能存在对公司成本估计偏低、毛利高估，从而导致对公司未来 3 年盈利预测值高于实际值的风险。

**AI 落地不及预期的风险：**公司 AI 产品已经实现商业化，随着 AI 产业快速发展，AI 产品需要迅速迭代以满足市场需求，若公司新一代 AI 产品研发不及预期，将影响公司未来产品的市场份额；同时，目前市场已有同类产品上市或在研竞品，未来商业化预计会面临激烈竞争，出现商业价值低或不及预期的风险，如果不能如期获得市场认可，将会对公司经营发展产生不利影响。

**产品价格下行及毛利率下降的风险：**公司产品迭代速度较快，通常上一代产品价格下行速度较快。未来，如果产品迭代速度放缓，旧产品价格持续下行，将对公司整体营收和毛利率产生不利影响。

**技术被赶超或替代的风险：**公司所处行业属于技术密集型行业，涉及软件平台构建、AI 产品研发等技术，在未来提升研发技术能力的竞争中，如果公司不能准确把握行业技术的发展趋势，在技术开发方向决策上发生失误；或研发项目未能顺利推进，未能及时将新技术运用于产品开发和升级，出现技术被赶超或替代的情况，公司将无法持续保持产品的竞争力，从而对公司的经营产生重大不利影响。

**宏观经济及行业波动风险：**如果未来宏观经济发生剧烈波动，导致终端需求下滑，将对公司的业务发展和经营业绩造成不利影响。

#### 相关研究报告：

《MINIMAX-WP (00100.HK) -Token 需求大幅增长，模型能力提升推动业务阶梯式增长》——2026-03-04  
《MINIMAX-WP (00100.HK) -MiniMax 2.5 让运行复杂 Agent 在经济上可行，Expert 功能沉淀具体领域 SOP》——2026-02-28  
《MINIMAX-WP (00100.HK) -全球化多模态大模型公司，高性价比构筑核心竞争力》——2026-02-13

## 财务预测与估值

资产负债表(百万元)					利润表(百万元)				
2025	2026E	2027E	2028E	2025	2026E	2027E	2028E		
现金及现金等价物	508	5566	5644	5598	营业收入	79	252	646	1293
应收款项	27	138	336	637	营业成本	59	170	399	727
存货净额	0	0	0	0	营业税金及附加	0	0	0	0
其他流动资产	20	65	166	333	销售费用	870	78	109	142
<b>流动资产合计</b>	<b>1007</b>	<b>6221</b>	<b>6622</b>	<b>7068</b>	管理费用	2034	411	553	698
固定资产	26	1	1	0	财务费用	0	50	(42)	(27)
无形资产及其他	2	3	3	3	投资收益	0	0	0	0
投资性房地产	1	1	1	1	资产减值及公允价值变动	(2142)	0	0	0
长期股权投资	0	0	0	0	其他收入	5901	293	293	293
<b>资产总计</b>	<b>1036</b>	<b>6226</b>	<b>6626</b>	<b>7072</b>	营业利润	876	(164)	(81)	45
短期借款及交易性金融负债	3633	1959	2403	2665	营业外净收支	(1590)	0	(200)	(200)
应付款项	58	172	403	733	<b>利润总额</b>	<b>(714)</b>	<b>(164)</b>	<b>(281)</b>	<b>(155)</b>
其他流动负债	43	10	15	23	所得税费用	30	(41)	0	0
<b>流动负债合计</b>	<b>3734</b>	<b>2140</b>	<b>2821</b>	<b>3421</b>	少数股东损益	12	3	5	3
长期借款及应付债券	0	0	0	0	<b>调整后净利润</b>	<b>(251)</b>	<b>(328)</b>	<b>(320)</b>	<b>(163)</b>
其他长期负债	3	3	3	3					
<b>长期负债合计</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>3</b>	现金流量表(百万元)				
<b>负债合计</b>	<b>3737</b>	<b>2143</b>	<b>2824</b>	<b>3424</b>	<b>净利润</b>	<b>(251)</b>	<b>(328)</b>	<b>(320)</b>	<b>(163)</b>
少数股东权益	0	3	8	10	资产减值准备	0	0	0	0
股东权益	(2700)	4080	3795	3637	折旧摊销	2	1	1	1
<b>负债和股东权益总计</b>	<b>1036</b>	<b>6226</b>	<b>6626</b>	<b>7072</b>	公允价值变动损失	2142	0	0	0
					财务费用	0	50	(42)	(27)
关键财务与估值指标					营运资本变动	10	(75)	(63)	(130)
每股收益	(2.41)	(0.40)	(0.91)	(0.50)	其它	(494)	205	40	8
每股红利	0.00	0.00	0.00	0.00	<b>经营活动现金流</b>	<b>1409</b>	<b>(198)</b>	<b>(343)</b>	<b>(284)</b>
每股净资产	(8.61)	13.01	12.10	11.60	资本开支	2139	0	0	0
ROIC	-344%	-9%	-7%	-4%	其它投资现金流	17	0	(23)	(24)
ROE	28%	-3%	-8%	-4%	<b>投资活动现金流</b>	<b>2156</b>	<b>0</b>	<b>(23)</b>	<b>(24)</b>
毛利率	25%	32%	38%	44%	权益性融资	0	6930	0	0
EBIT Margin	-3648%	-162%	-64%	-21%	负债净变化	0	0	0	0
EBITDA Margin	-3646%	-161%	-64%	-21%	支付股利、利息	0	0	0	0
收入增长	159%	219%	156%	100%	其它融资现金流	(3346)	(1674)	443	262
净利润增长率	0%	31%	-2%	-49%	<b>融资活动现金流</b>	<b>(3346)</b>	<b>5256</b>	<b>443</b>	<b>262</b>
资产负债率	361%	34%	43%	49%	<b>现金净变动</b>	<b>219</b>	<b>5058</b>	<b>78</b>	<b>(46)</b>
息率	([Red]12 0.3)	([Red]92 .0)	([Red]94 .2)	([Red]18 5.4)	货币资金的期初余额	289	508	5566	5644
P/E-调整后	(120.3)	(92.0)	(94.2)	(185.4)	货币资金的期末余额	508	5566	5644	5598
P/B	(87.4)	57.8	62.2	64.9	企业自由现金流	(855)	(482)	(477)	(404)
EV/EBITDA	(83.2)	(585.7)	(575.9)	(876.1)	权益自由现金流	(4201)	(2205)	8	(115)

资料来源: Wind、国信证券经济研究所预测

# 免责声明

## 分析师声明

作者保证报告所采用的数据均来自合规渠道；分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求独立、客观、公正，结论不受任何第三方的授意或影响；作者在过去、现在或未来未就其研究报告所提供的具体建议或所表述的意见直接或间接收取任何报酬，特此声明。

## 国信证券投资评级

投资评级标准	类别	级别	说明
报告中投资建议所涉及的评级（如有）分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 到 12 个月内的相对市场表现，也即报告发布日后的 6 到 12 个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A 股市场以沪深 300 指数（000300.SH）作为基准；新三板市场以三板成指（899001.CSI）为基准；香港市场以恒生指数（HSI.HI）作为基准；美国市场以标普 500 指数（SPX.GI）或纳斯达克指数（IXIC.GI）为基准。	股票 投资评级	优于大市	股价表现优于市场代表性指数 10%以上
		中性	股价表现介于市场代表性指数 ±10%之间
		弱于大市	股价表现弱于市场代表性指数 10%以上
		无评级	股价与市场代表性指数相比无明确观点
	行业 投资评级	优于大市	行业指数表现优于市场代表性指数 10%以上
		中性	行业指数表现介于市场代表性指数 ±10%之间
		弱于大市	行业指数表现弱于市场代表性指数 10%以上

## 重要声明

本报告由国信证券股份有限公司（已具备中国证监会许可的证券投资咨询业务资格）制作；报告版权归国信证券股份有限公司

关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以我公司向客户发布的本报告完整版本为准。

本报告基于已公开的资料或信息撰写，但我公司不保证该资料及信息的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映我公司于本报告公开发布当日的判断，在不同时期，我公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。我公司不保证本报告所含信息及资料处于最新状态；我公司可能随时补充、更新和修订有关信息及资料，投资者应当自行关注相关更新和修订内容。我公司或关联机构可能会持有本报告中所提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或金融产品等相关服务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中所提及的意见或建议不一致的投资决策。

本报告仅供参考之用，不构成出售或购买证券或其他投资标的的要约或邀请。在任何情况下，本报告中的信息和意见均不构成对任何个人的投资建议。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。投资者应结合自己的投资目标和财务状况自行判断是否采用本报告所载内容和信息并自行承担风险，我公司及雇员对投资者使用本报告及其内容而造成的一切后果不承担任何法律责任。

## 证券投资咨询业务的说明

本公司具备中国证监会核准的证券投资咨询业务资格。证券投资咨询，是指从事证券投资咨询业务的机构及其投资咨询人员以下列形式为证券投资人或者客户提供证券投资分析、预测或者建议等直接或者间接有偿咨询服务的活动：接受投资人或者客户委托，提供证券投资咨询服务；举办有关证券投资咨询的讲座、报告会、分析会等；在报刊上发表证券投资咨询的文章、评论、报告，以及通过电台、电视台等公众传播媒体提供证券投资咨询服务；通过电话、传真、电脑网络等电信设备系统，提供证券投资咨询服务；中国证监会认定的其他形式。

发布证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。

## 国信证券经济研究所

### 深圳

深圳市福田区福华一路 125 号国信金融大厦 36 层  
邮编：518046 总机：0755-82130833

### 上海

上海浦东民生路 1199 弄证大五道口广场 1 号楼 12 层  
邮编：200135

### 北京

北京西城区金融大街兴盛街 6 号国信证券 9 层  
邮编：100032