

2025年12月08日



DeepSeek 发布 V3.2 双模型，谷歌 DeepThink 引领推理革命

—计算机行业周报

推荐(维持)

分析师：任春阳 S1050521110006

reny@cfsc.com.cn

行业相对表现

表现	1M	3M	12M
计算机(申万)	-3.0	-3.1	7.2
沪深300	-1.2	3.5	16.3

市场表现



资料来源：Wind, 华鑫证券研究

相关研究

- 《计算机行业周报：ClaudeOpus4.5发布，DeepSeekMath-V2重磅开源》2025-12-01
- 《计算机行业点评报告：小马智行（PONY.0）：商业化里程碑确立，跨城深耕与全球布局双驱》2025-11-29
- 《计算机行业点评报告：DoorDash（DASH.0）：增长动能强劲，并购与投资驱动全球化新阶段》2025-11-29

投资要点

算力：算力租赁价格平稳，DeepSeek 发布 V3.2 双模型

2025年12月1日，DeepSeek 团队在 ChatGPT 发布三周年之际推出 DeepSeek-V3.2 系列开源大模型，以“双模型+三大技术突破”实现开源模型向第一梯队的逆袭，大幅缩小与闭源模型的性能差距。该系列包含两款定位差异化的模型：标准版 DeepSeek-V3.2 作为“日常全能手”，平衡推理能力与输出效率，推理水平达 GPT-5 级别，支持双模式工具调用，泛化能力强；增强版 Speciale 聚焦极致推理，在数学证明、国际竞赛等复杂任务上媲美 Gemini-3.0-Pro，仅面向科研场景开放测试。核心技术突破包括：DSA 稀疏注意力机制降低长序列计算成本，可扩展强化学习协议解锁高阶能力，大规模 Agent 任务合成补齐泛化短板。

AI 应用：Character.AI 周访问量环比+4.85%，谷歌 DeepThink 引领推理革命

2025 年 12 月 5 日，谷歌 DeepMind 正式推出了其新一代人工智能模型——Gemini3DeepThink，该模型在多项国际顶尖测试中表现出色，尤其在高难度数学、科学与逻辑推理任务上实现了显著突破。这一版本是基于前代 Gemini2.5DeepThink 的进一步升级，通过引入“并行思考”机制，模型能够同时探索多种假设与解决路径，从而在处理复杂问题时展现出更强大的推理与创造能力。其发布不仅标志着人工智能在深度推理领域迈出了关键一步，也被视为向通用人工智能（AGI）目标迈进的重要里程碑。

AI 融资动向：BlackForestLabs 完成 3 亿美元融资，实现技术架构代际升级

2025 年 12 月 2 日，德国开源图像生成独角兽 BlackForestLabs 宣布完成 3 亿美元 B 轮融资，估值飙升至 32.5 亿美元。本轮融資由 Salesforce Ventures 和知名风投 AnjneyMidha 领投，Nvidia、Canva、Figma Ventures、Temasek 以及多家风险投资机构参投。此次融资的核心优势在于其技术架构的代际升级。公司最新发布的 FLUX.2 模型系列摒弃了传统的纯扩散架构，转而采用了更为先进的“流匹配（FlowMatching）”技术。虽然两者都涉及从噪声中还原图像，但流匹配通过更高效的数学路径执行任务，大幅降低了

推理成本。更关键的是，FLUX.2 融合了 Transformer 架构元素——这种通常用于大语言模型（LLM）的设计，使得视觉模型开始具备更强的逻辑理解能力。配合引入 MistralAI 的开源视觉模型与潜在空间（LatentSpace）压缩技术，FLUX.2 成功在提升光影渲染逼真度的同时，实现了推理速度的数量级跃升。

■ 投资建议

2025 年 12 月 1 日，字节跳动豆包团队正式推出豆包手机助手技术预览版。该助手基于豆包大模型研发，现阶段在推理效率、视觉理解等核心技术领域的性能表现已跻身国际一流梯队。通过与手机厂商在操作系统层面开展深度协同合作，豆包手机助手构建了多元核心能力体系，涵盖语音唤醒、侧边键唤醒及耳机唤醒等多场景唤醒方式，同时具备跨界面交互、相册多模态修图、AI 驱动的跨应用比价下单与行程规划等复杂任务自动化处理能力，同步搭载记忆功能及 Pro 专业模式。该产品现阶段采用生态合作模式推进市场化落地，字节跳动豆包团队明确表示无自研手机的战略规划。当前，用户可通过字节跳动与中兴合作推出的 nubiaM153 工程样机体验该技术预览版，该机型定价 3499 元。官方同步提示，受大模型技术本身不确定性因素影响，演示场景无法确保 100% 复现，后续将持续投入资源推进产品迭代与技术优化。从行业发展视角来看，手机厂商与大模型厂商深度协同合作已成为当前行业发展的明确趋势，此次豆包手机助手的发布正是这一趋势下的重要实践探索。此次豆包手机助手的发布，为 AI 应用场景深化及端侧产业变革形成有力支撑。其复杂任务的高成功率，中兴合作工程机 nubiaM153 首销即售罄，跨应用比价下单、行程规划等场景落地，且以生态协同合作联动中兴并洽谈多厂商，印证了 AI 向端侧渗透的可行性与端侧智能市场高景气度。当前端侧 AI 商业化进程中，豆包依托端侧隐私保护、低延迟算力及多模态核心能力，加速 AI 在日常、办公等场景渗透，为 AI 应用提供优质端侧落地样本。此次发布有望为 AI 应用及端侧板块注入场景落地与商业化变现双重动能，我们持续看好相关板块。

中长期，建议关注加快扩张算力业务的精密零部件龙头迈信林（688685.SH）、新能源业务高增并供货科尔摩根等全球电机巨头的唯科科技（301196.SZ），AI 智能文字识别与商业大数据领域巨头的合合信息（688615.SH）、深耕工业 AI 与软件并长期服务高端装备等领域头部客户的能科科技（603859.SH）。

■ 风险提示

1) AI 底层技术迭代速度不及预期。2) 政策监管及版权风

险。3) AI应用落地效果不及预期。4) 推荐公司业绩不及预期风险。

重点关注公司及盈利预测

公司代码	名称	2025-12-08		EPS		PE		投资评级
		股价	2024	2025E	2026E	2024	2025E	
301196.SZ	唯科科技	77.51	1.76	2.53	3.34	44.04	30.64	23.21
603859.SH	能科科技	40.33	0.78	0.96	1.18	51.71	42.01	34.18
688615.SH	合合信息	197.18	4.01	3.37	4.11	49.17	58.51	47.98
688685.SH	迈信林	50.98	0.31	1.64	2.26	164.45	31.09	22.56

资料来源：Wind，华鑫证券研究

正文目录

1、算力动态：算力租赁价格平稳，DEEPSEEK 发布 V3.2 双模型	5
1.1、Tokens 跟踪.....	5
1.2、数据跟踪：算力租赁价格平稳	6
1.3、产业动态：DeepSeek 发布 V3.2 双模型，引领开源大模型推理与智能体能力新高度	6
2、AI 应用动态：CHARACTER.AI 周访问量环比+4.85%，谷歌 DEEPTHINK 引领推理革命	10
2.1、周流量跟踪：Character.AI 周访问量环比+4.85%.....	10
2.2、产业动态：谷歌 DeepThink 引领推理革命，人工智能竞争迈入新阶段	10
3、AI 融资动向：BLACKFORESTLABS 完成 3 亿美元融资，实现技术架构代际升级	13
4、行情复盘	15
5、投资建议	17
6、风险提示	17

图表目录

图表 1：TOKENS 规模 LEADERBOARD 示意.....	5
图表 2：市场份额占据示意	5
图表 3：上周算力租赁情况	6
图表 4：DEEPSEEK-V3.2 与其他模型在 AGENT 工具调用评测集的得分	6
图表 5：模型在推理与智能体能力任务中的表现对比	7
图表 6：DEEPSEEK-V3.2 注意力架构 DSA 示意图	7
图表 7：DEEPSEEK-V3.2-SFT 基于合成通用智能体数据的 RL 训练性能曲线	8
图表 8：DEEPSEEK-V3.2 与其他模型在多维度基准测试的得分	9
图表 7：2025.11.27-2025.12.3 AI 相关网站流量	10
图表 8：GEMINI3DEEPTHINK 在 HUMANITY' SLASTEXAM、GPQADIAMOND、ARC-AGI-2 中的表现	11
图表 9：GEMINI3DEEPTHINK 在 ARC-AGI-1 中的表现	11
图表 10：GEMINI 3PRO 网页端市占率	12
图表 11：GOOGLE GEMINI 每日访问量	12
图表 12：上周 AI 初创公司融资动态	13
图表 13：上周（12.1-12.5 日）指数日涨跌幅	15
图表 14：上周（12.1-12.5 日）AI 算力指数内部涨跌幅度排名	15
图表 15：上周（12.1-12.5 日）AI 应用指数内部涨跌幅度排名	16
图表 16：重点关注公司及盈利预测	17

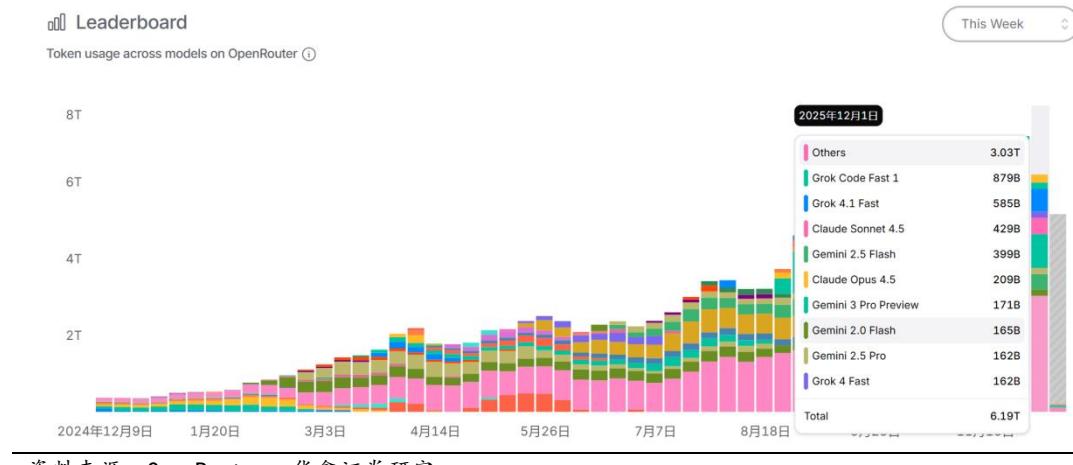
1、算力动态：算力租赁价格平稳，DeepSeek 发布 V3.2 双模型

1.1、Tokens 跟踪

根据 OpenRouter 公开数据，2025 年 12 月 1 日至 12 月 7 日，周度 token 消耗量有所下降，调用量为 6.19T，环比上周-13.9%。在 tokens 规模 leaderboard 前五名中，xAI 旗下 GrokCodeFast1 以 879Btokens 位居榜首；其另一产品 Grok4.1Fast 以 585Btokens 位列第二；Anthropic 的 ClaudeSonnet4.5 以 429Btokens 排第三；Google 的 Gemini2.5Flash 和 Anthropic 的 ClaudeOpus4.5 分别以 399B、209Btokens 位列第四、第五。

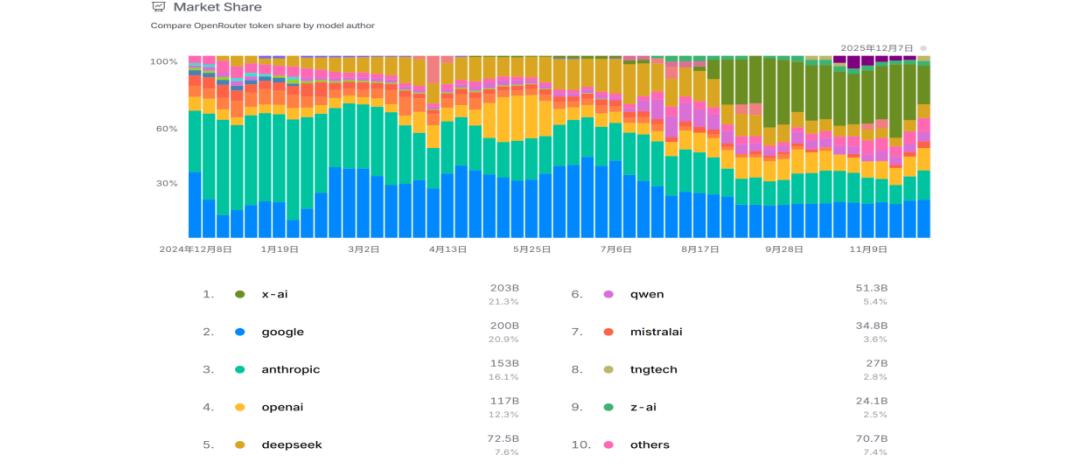
从市场份额维度来看，xAI 以 203Btokens 占据 21.3%的份额，稳居首位；Google 以 200Btokens 占比 20.9%，位列第二；Anthropic、OpenAI、Deepseek 则分别以 153B、117B、72.5Btokens，对应占据 16.1%、12.3%、7.6%的市场份额。

图表 1：tokens 规模 leaderboard 示意



资料来源：OpenRouter，华鑫证券研究

图表 2：市场份额占据示意



资料来源：OpenRouter，华鑫证券研究

请阅读最后一页重要免责声明

1.2、数据跟踪：算力租赁价格平稳

上周算力租赁价格平稳。具体来看，显卡配置为 A100-40G 中，腾讯云 16 核+96G 价格为 5.73 元/时，阿里云 12 核+94GiB 价格为 31.58 元/时；显卡配置为 A800-80G 中，恒源云 16 核+256G 价格为 7.50 元/时。

图表 3：上周算力租赁情况

显卡配置	CPU	内存	磁盘大小 (G)	平台名称	价格 (每小时)	价格环比上周
A100-40G	16	96	可自定，额外收费	腾讯云	5.73/元	0.00%
	12核	94G	可自定，额外收费	阿里云	31.58/元	0.00%
A100-80G	13	128	系统盘：20G 数据盘：50GB	恒源云	-	0.00%
A800-80G	16	256	系统盘：20G 数据盘：50GB	恒源云	7.50/元	0.00%

资料来源：腾讯云，阿里云，恒源云，华鑫证券研究

1.3、产业动态：DeepSeek 发布 V3.2 双模型，引领开源大模型推理与智能体能力新高度

2025 年 12 月 1 日，DeepSeek 团队在 ChatGPT 发布三周年之际推出的 DeepSeek-V3.2 系列，不仅填补了开源模型与闭源模型的性能鸿沟，更以“双模型+硬核技术”的组合，让开源大模型重新跻身全球第一梯队。

DeepSeek-V3.2 系列包含两个定位截然不同的模型，精准覆盖不同场景的需求。

作为聚焦实用的基础版本，DeepSeek-V3.2 的核心目标是平衡推理能力与输出效率——它既能胜任日常问答、通用 Agent 任务等真实场景需求，又通过优化大幅缩短了输出长度，减少用户等待时间。在性能上，它的推理能力达到 GPT-5 水平，略低于 Gemini-3.0-Pro；但相比 Kimi-K2-Thinking 等模型，输出长度大幅缩减，计算开销显著降低。更关键的是，它是 DeepSeek 首个“思考者+工具调用”双模模型：支持思考/非思考双模式工具调用，基于 1800+环境、85000+复杂指令的大规模 Agent 训练数据，泛化能力强。

作为 V3.2 的长思考增强版，Speciale 版本融合了 DeepSeek-Math-V2 的定理证明能力，在指令跟随、数学证明、逻辑验证等复杂任务上的表现媲美 Gemini-3.0-Pro，甚至在 IMO2025、CMO2025、ICPCWorldFinals2025 等国际顶级竞赛中斩获金。不过，这个版本并非为日常场景设计：它目前仅提供研究使用，不支持工具调用，也未针对日常对话与写作做专项优化；同时，它完成高复杂度任务时消耗的 Tokens 更多、成本更高，但在高度复杂任务上的表现大幅优于标准版。目前，Speciale 仅开放临时 API 供测试使用。

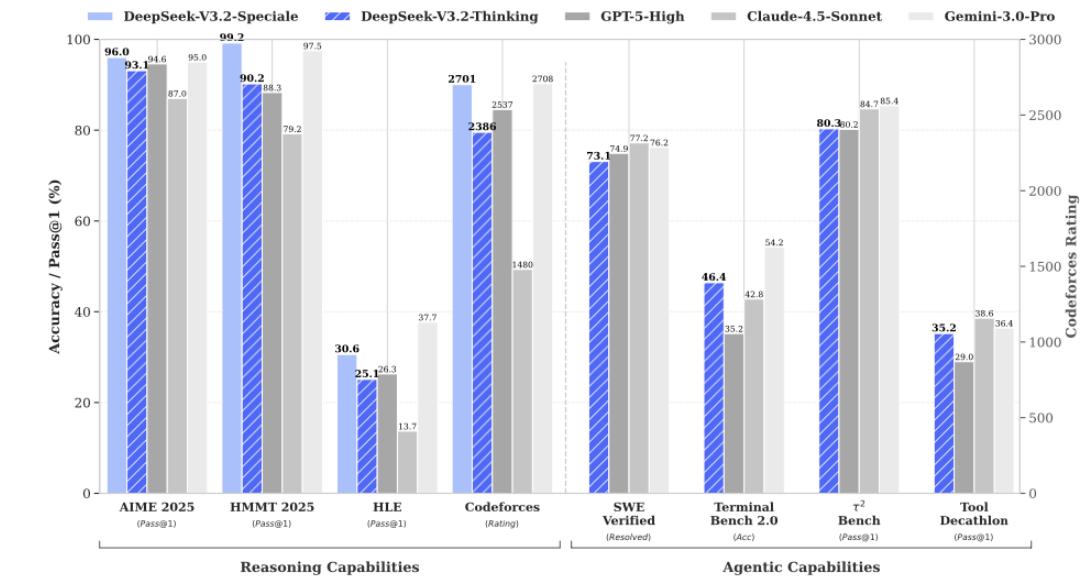
图表 4：DeepSeek-V3.2 与其他模型在 Agent 工具调用评测集的得分

	Benchmark	Claude-4.5-Sonnet	GPT-5-High	Gemini-3.0-Pro	Kimi-K2-Thinking	Minimax-M2	DeepSeek-V3.2-Thinking
ToolUse	t ² -Bench	84.7	80.2	85.4	74.3	76.9	80.3
	MCP-Universe	46.5	47.9	50.7	35.6	29.4	45.9
	MCP-Mark	33.3	50.9	43.1	20.4	24.4	38.0
	Tool-Decathlon	38.6	29.0	36.4	17.6	16.0	35.2

资料来源：量子位，华鑫证券研究

请阅读最后一页重要免责声明

图表 5：模型在推理与智能体能力任务中的表现对比



资料来源：TechWeb，华鑫证券研究

DeepSeek 团队精准识别了开源模型的三大短板：长序列效率低、训练资源不足、Agent 泛化能力弱，并通过三项硬核技术实现了突破。

突破一：DSA 稀疏注意力机制，解决长文本“计算负担”

传统注意力机制在处理长序列时计算复杂度为 $O(L^2)$ ，严重限制了模型的部署效率与训练扩展性。DeepSeek 团队的 DSA (DeepSeekSparseAttention) 稀疏注意力机制，将计算复杂度降至 $O(L \cdot k)$ (k 远小于 L)，同时让模型在长上下文任务中显著加速推理且无明显性能损失。

DSA 由“闪电索引器”和“细粒度 Token 选择”两个组件构成：前者快速计算新 Token 与历史 Token 的相关性分数，后者选择 Top-k 最相关 Token 进行注意力计算。团队通过两阶段训练策略（先训练索引器、再加入筛选机制），让 V3.2 在 128k 长度序列上的推理成本比 V3.1-Terminus 降低了好几倍——在 H800 集群测试中，当序列长度达 128k 时，预训练阶段每百万 Token 成本从 0.7 美元降至 0.2 美元，解码阶段从 2.4 美元降至 0.8 美元。

图表 6：DeepSeek-V3.2 注意力架构 DSA 示意图

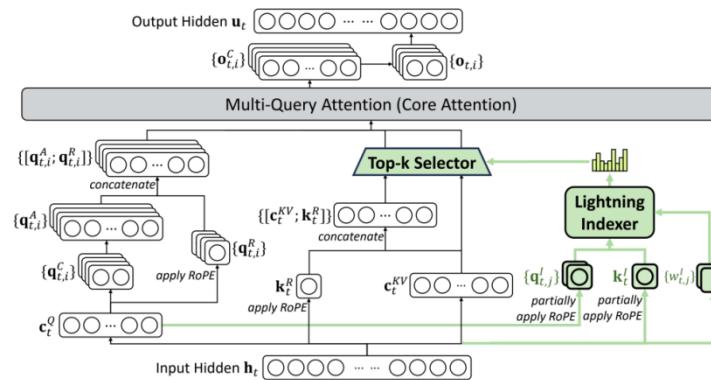


Figure 2 | Attention architecture of DeepSeek-V3.2, where DSA is instantiated under MLA. The green part illustrates how DSA selects the top-k key-value entries according to the indexer.

资料来源：量子位，华鑫证券研究

请阅读最后一页重要免责声明

突破二：可扩展强化学习协议，释放“高阶能力”

此前开源模型在训练阶段的计算资源投入不足，限制了其在困难任务上的表现。DeepSeek 团队开发了稳定且可扩展的强化学习（RL）协议，让后训练阶段的计算预算超过了预训练成本的 10%，从而解锁了模型的高级能力。

为了稳定扩展 RL 计算规模，团队做了三项优化：一是无偏 KL 估计，修正原始 K3 估计器的系统性误差；二是“离群序列掩码”策略，过滤掉偏离过大的负样本序列，避免干扰训练；三是 KeepRouting 操作，通过保存推理路径并在训练时强制使用相同路径，从而提升数据一致性。同时，团队采用“专家蒸馏”策略：先为数学、编程等任务训练专门的“专家模型”，再用这些模型生成的数据训练最终模型，进一步强化了复杂任务能力。

图表 7：DeepSeek-V3.2-SFT 基于合成通用智能体数据的 RL 训练性能曲线

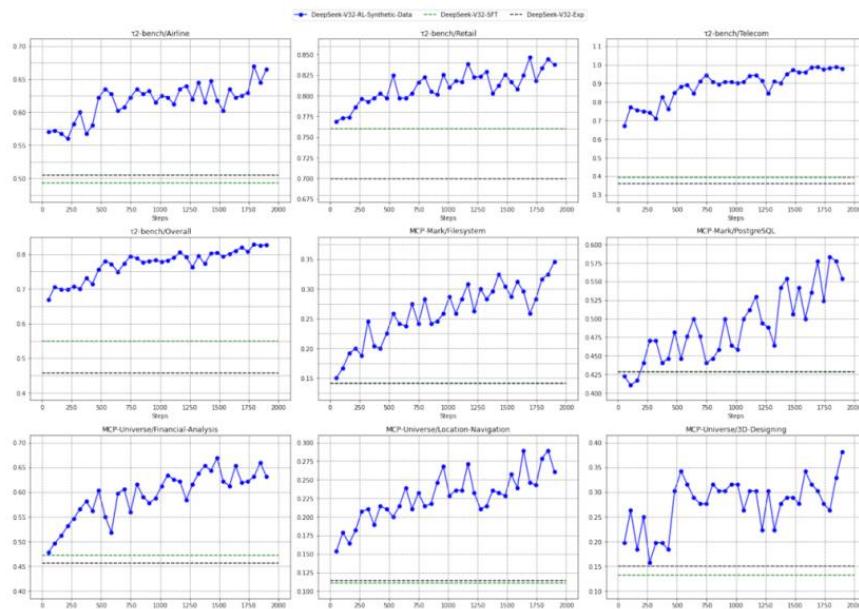


Figure 5 | RL training of DeepSeek-V3.2-SFT using exclusively synthetic general agent data.

资料来源：量子位，华鑫证券研究

突破三：大规模 Agent 任务合成，补齐“泛化短板”

DeepSeek 团队精心设计了一种全新的思考上下文管理机制：仅在引入新的用户消息时才会丢弃历史推理内容，而仅添加工具相关消息时，推理内容将得以保留。即使推理痕迹被删除，工具调用历史及其结果仍会存留在上下文中。在冷启动阶段，团队通过精心设计的系统提示，使模型能够在推理过程中自然地插入工具调用。此外，团队还开发了一个自动环境合成 pipeline，成功生成了 1827 个任务导向的环境和 85000 个复杂提示。

在代码 Agent 方面，团队从 GitHub 挖掘数百万个 issue-PR 对，搭建了数万个可执行的软件问题解决环境；搜索 Agent 则采用多 Agent pipeline 生成训练数据，覆盖长尾实体与复杂验证场景。

评测结果显示，DeepSeek-V3.2 在 SWE-Verified 上达到 73.1% 的解决率，在 TerminalBench2.0 上准确率达 46.4%，大幅超越现有开源模型；在 MCP-Universe 等工具使用基准测试中，性能已接近开源模型水平，这证明它能将推理策略泛化到训练时未见过的 Agent 场景。

图表 8: DeepSeek-V3.2 与其他模型在多维度基准测试的得分

Benchmark (Metric)		Claude-4.5-Sonnet	GPT-5 High	Gemini-3.0 Pro	Kimi-K2 Thinking	MiniMax M2	DeepSeek-V3.2 Thinking
English	MMLU-Pro (EM)	88.2	87.5	90.1	84.6	82.0	85.0
	GPQA Diamond (Pass@1)	83.4	85.7	91.9	84.5	77.7	82.4
	HLE (Pass@1)	13.7	26.3	37.7	23.9	12.5	25.1
Code	LiveCodeBench (Pass@1-COT)	64.0	84.5	90.7	82.6	83.0	83.3
	Codeforces (Rating)	1480	2537	2708	-	-	2386
Math	AIME 2025 (Pass@1)	87.0	94.6	95.0	94.5	78.3	93.1
	HMMT Feb 2025 (Pass@1)	79.2	88.3	97.5	89.4	-	92.5
	HMMT Nov 2025 (Pass@1)	81.7	89.2	93.3	89.2	-	90.2
	IMOAnswerBench (Pass@1)	-	76.0	83.3	78.6	-	78.3
Code Agent	Terminal Bench 2.0 (Acc)	42.8	35.2	54.2	35.7	30.0	46.4
	SWE Verified (Resolved)	77.2	74.9	76.2	71.3	69.4	73.1
	SWE Multilingual (Resolved)	68.0	55.3	-	61.1	56.5	70.2
Search Agent	BrowseComp (Pass@1)	24.1	54.9	-	-/60.2*	44.0	51.4/67.6*
	BrowseCompZh (Pass@1)	42.4	63.0	-	62.3	48.5	65.0
	HLE (Pass@1)	32.0	35.2	45.8	44.9	31.8	40.8
ToolUse	τ^2 -Bench(Pass@1)	84.7	80.2	85.4	74.3	76.9	80.3
	MCP-Universe (Success Rate)	46.5	47.9	50.7	35.6	29.4	45.9
	MCP-Mark (Pass@1)	33.3	50.9	43.1	20.4	24.4	38.0
	Tool-Decathlon (Pass@1)	38.6	29.0	36.4	17.6	16.0	35.2

资料来源：量子位，华鑫证券研究

 公众号 · 量子位

2、AI 应用动态：Character.AI 周访问量环比+4.85%，谷歌 DeepThink 引领推理革命

2.1、周流量跟踪：Character.AI 周访问量环比+4.85%

本期（2025.11.27-2025.12.3）AI 相关网站流量数据：访问量前三位分别为 ChatGPT（1332.0M）、Bing（772.5M）和 Gemini（374.6M），访问量环比增速第一为 Character.AI（4.85%）；平均停留时长前三位分别为 Character.AI（00:18:30）、Discord（00:10:49）和 Kimi（00:08:36）；平均停留时长环比增速第一为 Kimi（1.57%）。

图表 7：2025.11.27-2025.12.3 AI 相关网站流量

应用	应用类型	归属公司	周平均访问量 (M)	访问量环比	平均停留时长	时长环比
ChatGPT	聊天机器人	OpenAI	1332.0	-0.67%	6:33	0.26%
Bing	搜索	微软	772.5	-0.73%	7:56	1.06%
Gemini	聊天机器人	谷歌	374.6	2.80%	7:01	0.72%
Canva	在线设计	Canva	231.7	-0.26%	6:07	-0.81%
Discord	游戏社区	微软	131.4	-0.53%	10:49	-0.76%
Github	代码托管	微软	113.6	0.26%	6:29	-0.26%
Perplexity	AI 搜索	Perplexity	43.24	-1.32%	4:34	0.00%
Character.AI	聊天机器人	Character.AI	42.98	4.85%	18:30	0.82%
NotionAI	文本/笔记	Notion	38.43	1.53%	8:02	0.42%
DeepL	翻译工具	DeepL	29.15	-0.68%	2:33	0.00%
QuillBot	释义工具	QuillBot	13.98	1.38%	3:09	0.53%
Kimi	聊天机器人	MoonshotAI	9.78	-3.77%	8:36	1.57%
文心一言	聊天机器人	百度	0.92	2.96%	2:58	-2.73%

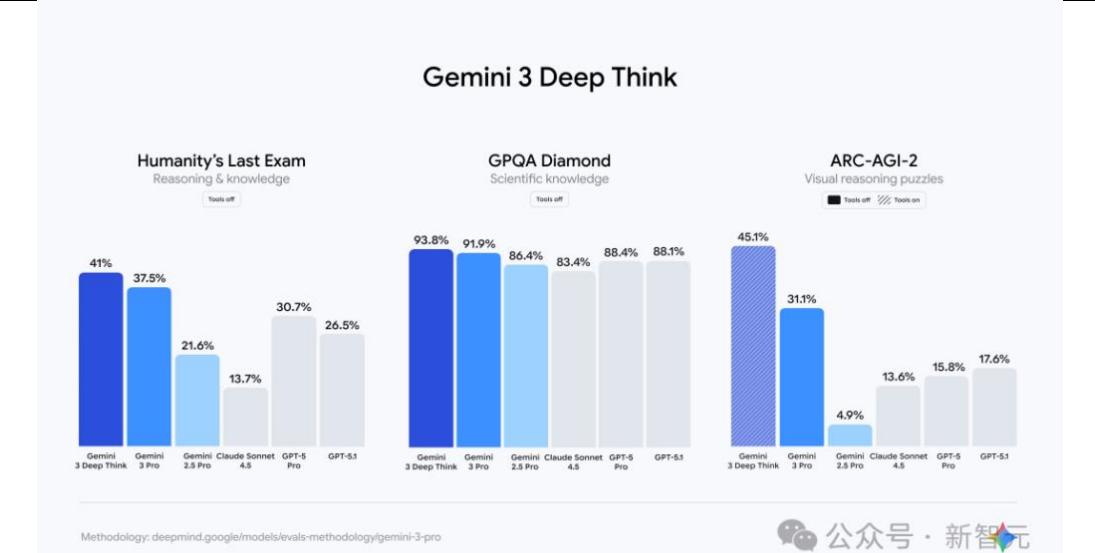
资料来源：similarweb，华鑫证券研究

2.2、产业动态：谷歌 DeepThink 引领推理革命，人工智能竞争迈入新阶段

2025 年 12 月 5 日，谷歌 DeepMind 正式推出了其新一代人工智能模型——Gemini3DeepThink，该模型在多项国际顶尖测试中表现出色，尤其在高难度数学、科学与逻辑推理任务上实现了显著突破。这一版本是基于前代 Gemini2.5DeepThink 的进一步升级，通过引入“并行思考”机制，模型能够同时探索多种假设与解决路径，从而在处理复杂问题时展现出更强大的推理与创造能力。其发布不仅标志着人工智能在深度推理领域迈出了关键一步，也被视为向通用人工智能（AGI）目标迈进的重要里程碑。

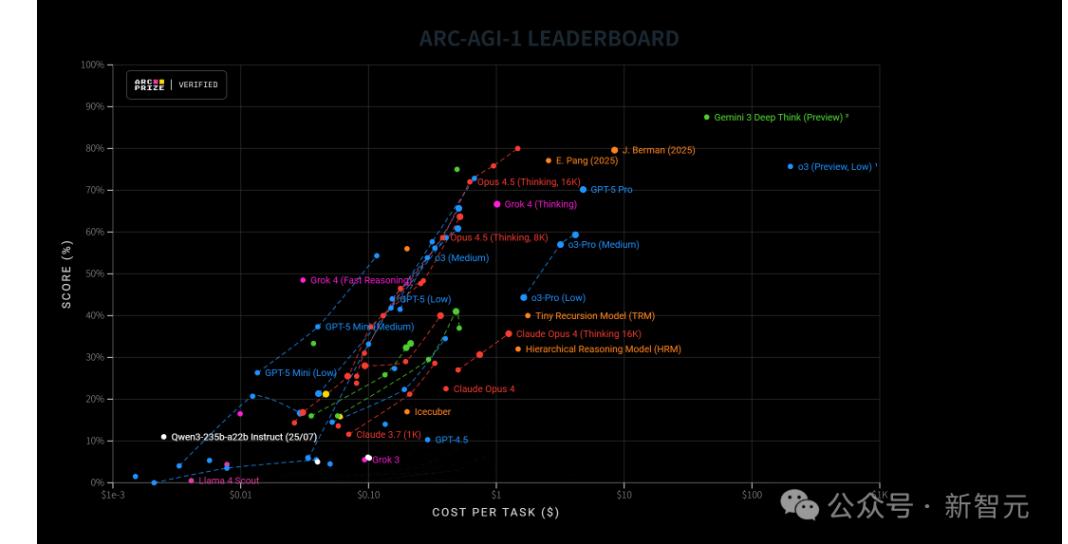
在性能方面，Gemini3DeepThink 在多个权威基准测试中取得了领先行业的成绩。在无需外部工具辅助的“Humanity’s LastExam”测试中，该模型获得了 41% 的高分；而在需要代码执行配合的 ARC-AGI-2 评估中，更是以 45.1% 的成绩刷新了全球纪录。相比之下，Gemini3Pro 版本在相同任务上的表现则明显逊色，尤其在涉及多步骤推理、物理模拟与创造性生成的场景中，DeepThink 版本展现出了极强的理解与执行能力。例如，在根据二维草图生成精确三维交互场景的任务中，DeepThink 版本能够高度还原原图细节，并在光影效果、物理交互等方面实现符合真实逻辑的模拟，体现出其在多模态理解与生成方面的显著优势。

图表 8：Gemini3DeepThink 在 Humanity’s LastExam、GPQADiamond、ARC-AGI-2 中的表现



资料来源：新智元，华鑫证券研究

图表 9：Gemini3DeepThink 在 ARC-AGI-1 中的表现



值得关注的是，推动这一技术突破的背后是一支由华人科学家领衔的精英研发团队。谷歌 DeepMind 近日宣布，将在新加坡组建一个专注于高级推理、LLM/RL，以及推进 Gemini、GeminiDeepThink 等最前沿 SOTA 模型的发展的新团队，该团队将由知名华人研究员 Yi Tay 负责领导，并向位于美国山景城的谷歌资深研究员 QuocLe 汇报。Yi Tay 强调人才密度的重要性，表示会以小而精的配置开展前沿攻关，并与多位重量级研究员保持紧密协作，共同推进 Gemini 系列模型的发展。

除了技术层面的突破，Gemini3 系列的发布也在市场层面引起了显著反响。自 Gemini3Pro 上线以来，谷歌 Gemini 系列产品的网页端访问量实现了显著增长，总体市场份额突破 15%，月访问量达到 13.51 亿次，环比增幅超过 14%。与此同时，尽管 OpenAI 旗下的 ChatGPT 在绝对访问量上仍保持领先，但其市场份额呈现持续下滑趋势，月访问量在近期首次跌破 60 亿次。另一方面，Grok 模型也凭借其 4.1 版本的更新实现了连续两个月的流量增长，访问量创下历史新高。这一动态反映出当前人工智能助手市场竞争日趋激烈。

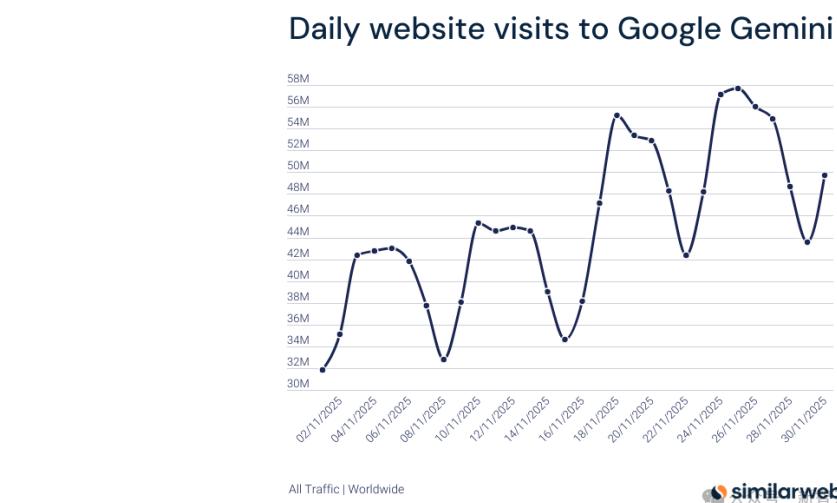
总体而言，Gemini3DeepThink 的推出，标志着人工智能竞赛的重点正从追求规模转向攻坚深层能力。其推理、问题解决及创造力的提升，是技术向具备逻辑、规划与常识理解的深层智能演进的关键一步。

图表 10: Gemini 3Pro 网页端市占率

	ChatGPT	Gemini	DeepSeek	Grok	Perplexity	Claude	Copilot
12个月前	87.0%	5.7%	/	/	2.1%	1.7%	1.6%
6个月前	79.8%	6.6%	5.7%	2.3%	1.7%	1.2%	1.2%
3个月前	78.1%	9.1%	4.0%	2.3%	1.9%	1.8%	1.2%
1个月前	72.9%	13.3%	4.1%	2.4%	%2.4	2.2%	1.2%
目前	71.3%	15.1%	4.0%	2.7%	2.2%	2.0%	1.2%

资料来源：新智元，华鑫证券研究

图表 11: GoogleGemini 每日访问量



资料来源：新智元，华鑫证券研究

3、AI 融资动向：BlackForestLabs 完成 3 亿美元融资，实现技术架构代际升级

2025 年 12 月 2 日，德国开源图像生成独角兽 BlackForestLabs 宣布完成 3 亿美元 B 轮融资，估值 32.5 亿美元。本轮融資由 Salesforce Ventures 和知名风投 AnjneyMidha 领投，Nvidia、Canva、Figma Ventures、Temasek 以及多家风险投资机构参投。

BlackForestLabs 于 2024 年正式成立，由机器学习领域研究人员 Robin Rombach、Patrick Esser 与 Andreas Blattmann 联合创办。该团队核心成员此前曾参与开源图像生成模型系列 StableDiffusion 的研发工作。

BlackForestLabs 本次融资的核心优势在于其技术架构的代际升级。公司最新发布的 FLUX. 2 模型系列摒弃了传统的纯扩散架构，转而采用更为先进的“FlowMatching”技术。尽管二者均涉及从噪声中还原图像，但流匹配技术通过更高效的数学路径执行任务，进一步降低了推理成本。更为关键的是，FLUX. 2 融合了 Transformer 架构元素——该架构通常应用于大型语言模型，使视觉模型具备了更强的逻辑理解能力。此外，FLUX. 2 还引入了 MistralAI 的开源视觉模型与潜在空间压缩技术，在提升光影渲染逼真度的同时，实现了推理速度的数量级提升。

在商业模式探索方面，BlackForestLabs 正实施一套成熟的“开源核心+专有增值”策略。该公司在 GitHub 上开放了 FLUX. 2 的基础版本，凭借极低的入门门槛，通过开发者社区构建起生态护城河；同时，针对企业级需求，推出了 FLUX. 2[pro]（极致画质）和 FLUX. 2[flex]（极致速度）两款专有版本。这一策略不仅维持了技术社区的活跃度和迭代效率，还为对性能和版权有严格要求的 B 端客户提供了付费选择。此前发布的 FLUX. 1 系列已累积数百万用户，充分证明了该模式在与 Midjourney 等闭源竞品竞争时的强大竞争力。

然而，BlackForestLabs 的战略布局远不止于文生图领域。公司明确表示，新注入的资金将用于构建集视觉感知、生成、记忆和推理于一体的超级模型，这标志着公司正从单一的 AIGC 内容生成向视觉智能体方向转型。通过赋予算法强大的推理能力，BlackForestLabs 实质上是在为机器人市场打造“感知系统”与“决策中枢”。未来，机器人不仅需高效处理传感器数据，更需基于视觉信息做出复杂的导航与操作决策。这一战略方向的转变，将使其直接与 LumaAI 等获得巨额融资的 AI 公司展开竞争，争夺物理世界 AI 的市场入场权。

图表 12：上周 AI 初创公司融资动态

应用	应用类型	领投方	融资轮	融资额	目前累计融资额	目前估值
BlackForestLabs	开源图像生成	Salesforce Ventures、AnjneyMidha	B 轮	3 亿美元	4.31 亿美元	32.5 亿美元

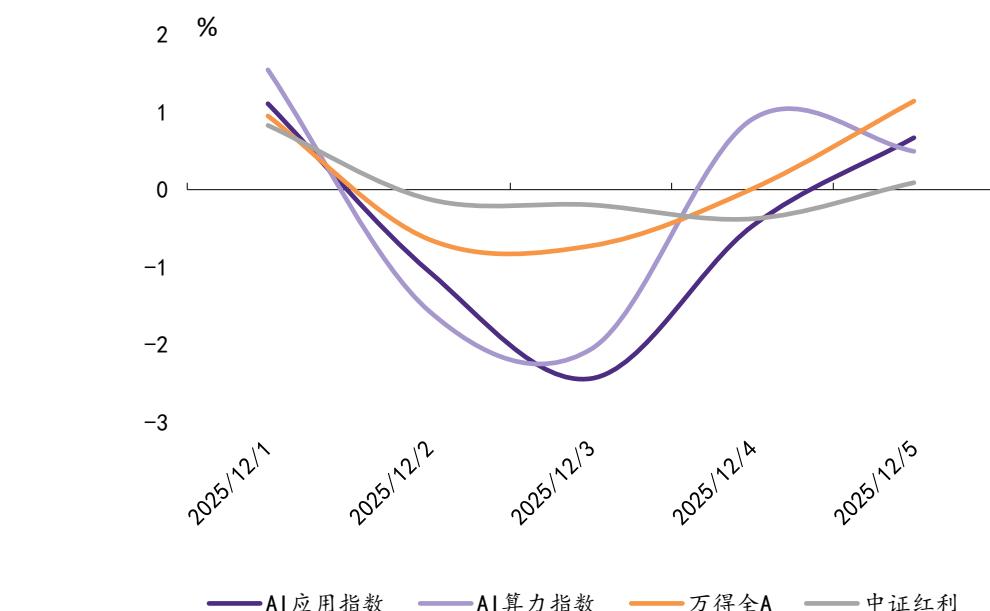
AngleHealth	AI 医疗福利平台	Portage	B 轮	1.34 亿美 元	1.96 亿美 元	-
Nevis	AI 原生财富管理平台	SequoiaCapital、ICONIQ 和 RibbitCapital	A 轮	3500 万美 元	4000 万美 元	-

资料来源：SaaSverse, wind, , 企查查, 华鑫证券研究

4、行情复盘

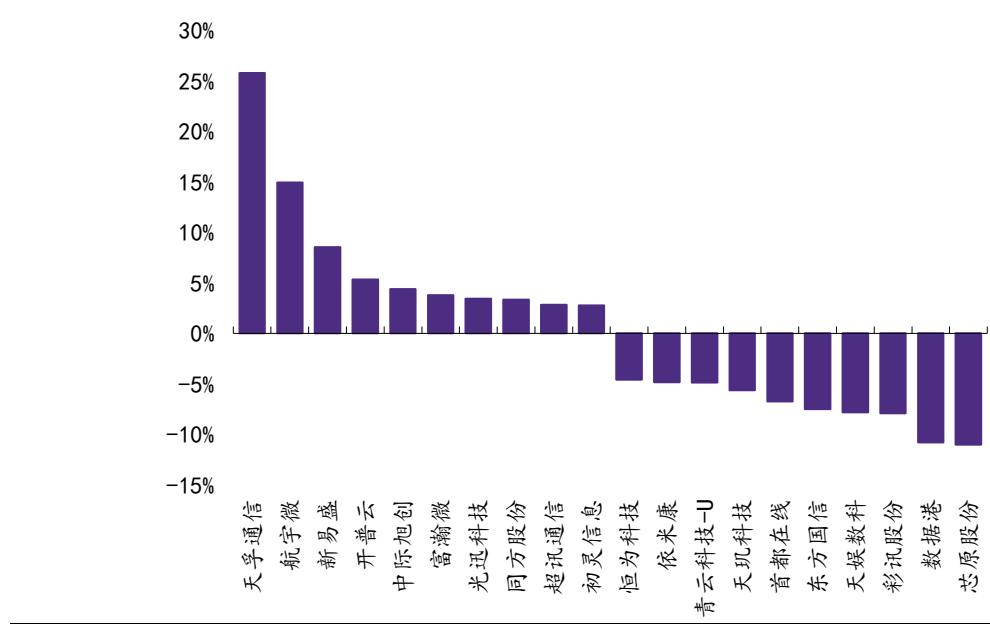
上周（12.1-12.5 日），AI 应用指数/AI 算力指数/万得全 A/中证红利日涨幅最大值分别为 1.11%/1.55%/1.14%/0.83%，AI 应用指数/AI 算力指数/万得全 A/中证红利日跌幅最大值分别为 -2.44%/-2.06%/-0.72%/-0.38%。AI 算力指数内部，天孚通信以+25.79%录得上周最大涨幅，芯原股份以-11.03%录得上周最大跌幅。AI 应用指数内部，睿创微纳以+10.84%录得上周最大涨幅，蓝色光标以-13.98%录得上周最大跌幅。

图表 13：上周（12.1-12.5 日）指数日涨跌幅



资料来源：wind, 华鑫证券研究

图表 14：上周（12.1-12.5 日）AI 算力指数内部涨跌幅度排名



资料来源：wind, 华鑫证券研究

请阅读最后一页重要免责声明

图表 15：上周（12.1-12.5 日）AI 应用指数内部涨跌幅度排名



资料来源：wind, 华鑫证券研究

5、投资建议

2025年12月1日，字节跳动豆包团队正式推出豆包手机助手技术预览版。该助手基于豆包大模型研发，现阶段在推理效率、视觉理解等核心技术领域的性能表现已跻身国际一流梯队。通过与手机厂商在操作系统层面开展深度协同合作，豆包手机助手构建了多元核心能力体系，涵盖语音唤醒、侧边键唤醒及耳机唤醒等多场景唤醒方式，同时具备跨界面交互、相册多模态修图、AI驱动的跨应用比价下单与行程规划等复杂任务自动化处理能力，同步搭载记忆功能及Pro专业模式。该产品现阶段采用生态合作模式推进市场化落地，字节跳动豆包团队明确表示无自研手机的战略规划。当前，用户可通过字节跳动与中兴合作推出的nubiaM153工程样机体验该技术预览版，该机型定价3499元。官方同步提示，受大模型技术本身不确定性因素影响，演示场景无法确保100%复现，后续将持续投入资源推进产品迭代与技术优化。从行业发展视角来看，手机厂商与大模型厂商深度协同合作已成为当前行业发展的明确趋势，此次豆包手机助手的发布正是这一趋势下的重要实践探索。此次豆包手机助手的发布，为AI应用场景深化及端侧产业变革形成有力支撑。其复杂任务的高成功率，中兴合作工程机nubiaM153首销即售罄，跨应用比价下单、行程规划等场景落地，且以生态协同合作联动中兴并洽谈多厂商，印证了AI向端侧渗透的可行性与端侧智能市场高景气度。当前端侧AI商业化进程中，豆包依托端侧隐私保护、低延迟算力及多模态核心能力，加速AI在日常、办公等场景渗透，为AI应用提供优质端侧落地样本。此次发布有望为AI应用及端侧板块注入场景落地与商业化变现双重动能，我们持续看好相关板块。

中长期，建议关注加快扩张算力业务的精密零部件龙头迈信林（688685.SH）、新能源业务高增并供货科尔摩根等全球电机巨头的唯科科技（301196.SZ），AI智能文字识别与商业大数据领域巨头的合合信息（688615.SH）、深耕工业AI与软件并长期服务高端装备等领域头部客户的能科科技（603859.SH）。

图表 16：重点关注公司及盈利预测

公司代码	名称	2025-12-08		EPS		PE		投资评级
		股价	2024	2025E	2026E	2024	2025E	
301196.SZ	唯科科技	77.51	1.76	2.53	3.34	44.04	30.64	23.21 买入
603859.SH	能科科技	40.33	0.78	0.96	1.18	51.71	42.01	34.18 买入
688615.SH	合合信息	197.18	4.01	3.37	4.11	49.17	58.51	47.98 买入
688685.SH	迈信林	50.98	0.31	1.64	2.26	164.45	31.09	22.56 买入

资料来源：Wind，华鑫证券研究

6、风险提示

1) AI底层技术迭代速度不及预期。2) 政策监管及版权风险。3) AI应用落地效果不及预期。4) 推荐公司业绩不及预期风险。

■ 中小盘&主题&北交所组介绍

任春阳：华东师范大学经济学硕士，6年证券行业经验，2021年11月加盟华鑫证券研究所，从事计算机与中小盘行业上市公司研究

周文龙：澳大利亚莫纳什大学金融硕士

陶欣怡：毕业于上海交通大学，于2023年10月加入团队。

倪汇康：金融学士，2025年8月加盟华鑫证券研究所。

■ 证券分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映了本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接受到任何形式的补偿。

■ 证券投资评级说明

股票投资评级说明：

	投资建议	预测个股相对同期证券市场代表性指数涨幅
1	买入	>20%
2	增持	10%—20%
3	中性	-10%—10%
4	卖出	<-10%

行业投资评级说明：

	投资建议	行业指数相对同期证券市场代表性指数涨幅
1	推荐	>10%
2	中性	-10%—10%
3	回避	<-10%

以报告日后的12个月内，预测个股或行业指数相对于相关证券市场主要指数的涨跌幅为标准。

相关证券市场代表性指数说明：A股市场以沪深300指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以恒生指数为基准；美国市场以道琼斯指数为基准。

■ 免责条款

请阅读最后一页重要免责声明

华鑫证券有限责任公司（以下简称“华鑫证券”）具有中国证监会核准的证券投资咨询业务资格。本报告由华鑫证券制作，仅供华鑫证券的客户使用。本公司不会因接收人收到本报告而视其为客户。

本报告中的信息均来源于公开资料，华鑫证券研究部门及相关研究人员力求准确可靠，但对这些信息的准确性及完整性不作任何保证。我们已力求报告内容客观、公正，但报告中的信息与所表达的观点不构成所述证券买卖的出价或询价的依据，该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。投资者应当对本报告中的信息和意见进行独立评估，并应同时结合各自的投资目的、财务状况和特定需求，必要时就财务、法律、商业、税收等方面咨询专业顾问的意见。对依据或者使用本报告所造成的一切后果，华鑫证券及/或其关联人员均不承担任何法律责任。本公司或关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露。

本报告中的资料、意见、预测均只反映报告初次发布时的判断，可能会随时调整。该等意见、评估及预测无需通知即可随时更改。在不同时期，华鑫证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。华鑫证券没有将此意见及建议向报告所有接收者进行更新的义务。

本报告版权仅为华鑫证券所有，未经华鑫证券书面授权，任何机构和个人不得以任何形式刊载、翻版、复制、发布、转发或引用本报告的任何部分。若华鑫证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，华鑫证券对此等行为不承担任何责任。本报告同时不构成华鑫证券向发送本报告的机构之客户提供的投资建议。如未经华鑫证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。华鑫证券将保留随时追究其法律责任的权利。请投资者慎重使用未经授权刊载或者转发的华鑫证券研究报告。

报告编号：HX-251208162023

请阅读最后一页重要免责声明

19