



DeepSeek 火爆全球，AI 即将迎来新一轮繁荣期

2025 年 2 月 12 日

概览【2.5-2.11】

- **股指动态：**2 月 5 日至 11 日，美股科技小幅回调，A 股计算机累计涨 15.65%。标普 500 指数累计+0.51%，纳斯达克综合指数-0.05%，费城半导体指数+2.08%；TAMAMA 科技指数-1.05%；纳斯达克中国金龙指数+2.63%；恒生科技指数+3.31%；计算机+15.65%。
- **个股表现：**2 月 5 日至 11 日，热门科技股大部分上涨。据统计，相比 2 月 5 日收盘价，2 月 11 日盘后，苹果合计+0.17%，英伟达+6.38%，特斯拉-13.13%，谷歌-3.14%，亚马逊-1.44%，META+2.12%，微软-0.45%，ARM-8.96%，英特尔+6.72%，高通-2.20%，AMD-0.81%。
- **10 年期国债及汇率：**2 月 11 日，美国 10 年期国债利率为 4.54%，相较 2 月 5 日累计上升 11bps；2 月 11 日，美元兑人民币中间价报 7.17；较 2 月 5 日价累计调贬 23 个基点。

● **核心观点：**

DeepSeek 火爆全球并影响全球人工智能产业链格局，以 AI Agent 为代表的 AI 应用及端侧 AI 有望迎来爆发期。2 月 9 日，沙特阿美旗下 Aramco Digital 宣布，DeepSeek 在其位于沙特阿拉伯达曼的数据中心开始运营。而此前 2 月 6 日，中国移动、中国联通、中国电信联合宣布已完成 DeepSeek - R1 及 V3 模型在云计算平台的全面接入。DeepSeek AI 助手一经推出已成为全球增速最快的 AI 应用之一，目前日活用户量已经突破 3000 万。

美国科技巨头持续投入基础设施算力，凸显科技巨头对人工智能领域前景的乐观态度。微软、亚马逊、谷歌、Meta 这四家美国科技公司 2024 年资本支出均达到史上最高点，预计 2025 年资本支出仍将高速增长。2 月 7 日亚马逊发布第四季度财报，公司业绩电话会议上表示计划在今年投资高达 1000 亿美元大部分用于数据中心建设。谷歌母公司 Alphabet 计划在 2025 年投入约 750 亿美元用于资本支出；Meta 预计 2025 年资本支出为 600 亿-650 亿美元。

我们认为：DeepSeek 通过架构优化和训练方法革新降低单次训练和推理的算力需求，压缩成本至行业极低水平，但并未打破 Scaling Laws 的本质与创新边界，算力需求依然长期增长，由于低成本将大幅推动应用普及，未来推理算力将接棒训练算力成为算力消耗主力，全球 AI+行业应用、端侧 AI 预计将迎来新一轮繁荣期。

- **风险提示：**技术迭代不及预期风险；科技巨头竞争加剧风险；法律监管风险；供应链风险；下游需求不及预期风险。

计算机行业

推荐 维持评级

分析师

吴砚靖

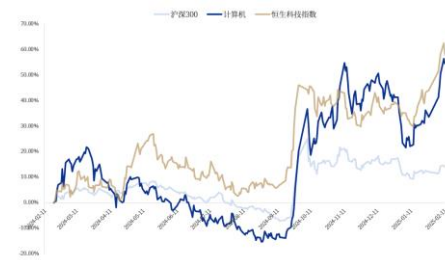
☎：010-66568589

✉：wuyanqing@chinastock.com.cn

分析师登记编码：S0130519070001

国内表现

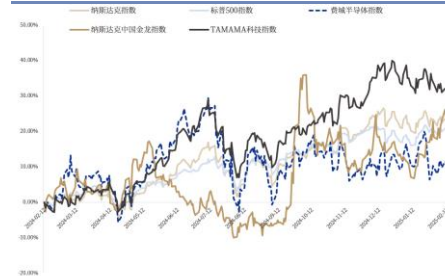
2025-02-11



资料来源：中国银河证券研究院

全球行情

2025-02-11



资料来源：中国银河证券研究院

相关研究

【银河计算机】全球科技动态追踪_降息改善流动性，重视港股科技投资机会

【银河计算机】全球科技动态追踪_特斯拉、谷歌等科技龙头陆续披露三季报

【银河计算机】全球科技动态追踪_大选对科技的影响：特朗普与哈里斯的不同路径

目录 Catalog

一、全球市场表现	3
(一) 股市动态	3
(二) 债市及汇率情况	3
(三) 重点科技股表现	3
二、行业要闻	5
(一) 算力及终端	5
(二) 大模型及云应用	5
三、风险提示	7

一、全球市场表现

(一) 股市动态

2月5日至11日，美股科技小幅回调，A股计算机累计涨15.65%。标普500指数累计+0.51%，纳斯达克综合指数-0.05%，费城半导体指数+2.08%；TAMAMA科技指数-1.05%；纳斯达克中国金龙指数+2.63%；恒生科技指数+3.31%；计算机+15.65%。

表1: 主要股指周变动

指数代码	指数简称	涨跌幅%				市盈率 PE (TTM)
		区间[02.05-02.11]	本月	本年度	2024	
SPX.GI	标普500指数	0.51	0.46	3.18	23.31	28.08
IXIC.GI	纳斯达克指数	-0.05	0.08	1.72	28.64	44.75
SOX.GI	费城半导体指数	2.08	1.28	2.01	19.27	53.12
8884057.WI	TAMAMA科技指数	-1.05	-1.25	-1.38	46.87	36.21
HXC.GI	纳斯达克中国金龙指数	2.63	4.79	9.28	4.43	25.46
HSTECH.HI	恒生科技指数	3.31	8.86	15.09	18.70	22.89
CI005027.WI	计算机	15.65	15.65	14.03	9.91	126.43

资料来源: WIND, 中国银河证券研究院

(二) 债市及汇率情况

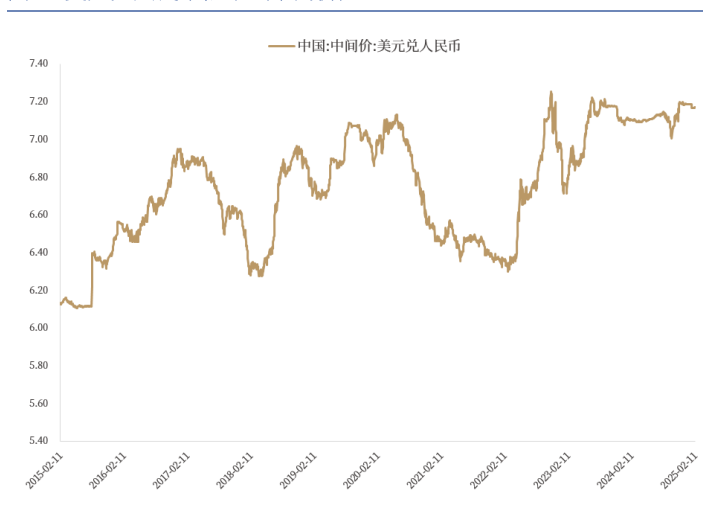
2月11日，美国10年期国债利率为4.54%，相较2月5日累计上升11bps；2月11日，美元兑人民币中间价报7.17；较2月5日价累计调贬23个基点。

图1: 美债收益率 (10年期)



资料来源: WIND, 中国银河证券研究院

图2: 美元兑人民币汇率 (中间价)



资料来源: WIND, 中国银河证券研究院

(三) 重点科技股表现

2月5日至11日，热门科技股大部分上涨。据统计，相比2月5日收盘价，2月11日盘

后，苹果合计+0.17%，英伟达+6.38%，特斯拉-13.13%，谷歌-3.14%，亚马逊-1.44%，META+2.12%，微软-0.45%，ARM-8.96%，英特尔+6.72%，高通-2.20%，AMD-0.81%。

美国科技巨头持续投入基础设施算力，凸显科技巨头对人工智能领域前景的乐观态度。微软、亚马逊、谷歌、Meta 这四家美国科技公司 2024 年资本支出均达到史上最高点，预计 2025 年资本支出仍将高速增长。2 月 7 日亚马逊发布第四季度财报，公司业绩电话会议上表示计划在今年投资高达 1000 亿美元大部分用于数据中心建设。谷歌母公司 Alphabet 计划在 2025 年投入约 750 亿美元用于资本支出；Meta 预计 2025 年资本支出为 600 亿-650 亿美元。

表2：重点公司周数据

所属板块	地区	证券代码	公司名称	股价 (美元/港元/新台币)		区间涨跌幅 (%)	总市值 (亿美元/亿港元/亿新台币)	[2025-02-11]	市销率 PS (TTM)	市盈率 PE (TTM)	PE(2023)	PE(2024)	PE(2025E)
				[2025-02-05]	[2025-02-11]								
算力及终端	美股	AAPL.O	苹果(Apple)	232.47	232.62	0.17	34,944.35	8.83	36.34	30.87	40.18	31.49	
		NVDA.O	英伟达(NVIDIA)	124.83	132.80	6.38	32,522.72	28.71	51.56	280.04	110.51	46.55	
		TSLA.O	特斯拉(TESLA)	378.17	328.50	-13.13	10,566.26	10.82	149.01	52.67	183.15	101.62	
		HPQ.N	惠普(HP)	32.57	33.15	1.78	310.88	0.58	11.20	9.14	11.03	11.67	
		CSCO.O	思科(CISCO)	62.57	62.43	-0.22	2,486.44	4.69	26.47	16.28	22.85	0.00	
		ASML.O	阿斯麦	742.59	752.98	1.62	2,965.47	10.02	37.41	34.34	33.90	29.32	
		AMD.O	超威半导体(AMD)	112.01	111.10	-0.81	1,800.35	6.98	109.71	278.85	119.39	39.24	
		INTC.O	英特尔(INTEL)	19.65	20.97	6.72	908.00	1.71	-4.84	125.43	-4.63	38.39	
		QCOM.O	高通(QUALCOMM)	175.86	171.99	-2.20	1,902.21	4.67	18.02	22.36	16.75	17.18	
		ARM.O	ARM	173.26	157.74	-8.96	1,662.58	45.01	206.28	147.03	424.91	325.44	
	ON.O	安森美半导体(ON SEMICONDUCTOR)	51.72	47.93	-7.33	201.99	2.85	12.84	16.47	16.96	15.04		
	0909.HK	明源云	2.91	3.45	18.56	67.06	3.79	-16.21	-8.66	0.00	-49.17		
	9698.HK	万国数据-SW	24.75	30.00	21.21	457.33	3.76	-10.35	-2.92	0.00	-60.96		
	1686.HK	新意向集团	3.84	4.46	16.15	104.32	3.89	11.50	7.98	11.01	15.11		
2330.TW	台积电	1,110.00	1,110.00	0.00	348,789.80	9.95	24.53	18.05	23.76	18.82			
2454.TW	联发科	1,525.00	1,510.00	-0.98	24,185.10	4.64	22.34	21.09	0.00	0.00			
互联网	美股	GOOGL.O	谷歌(ALPHABET)-A	191.33	185.32	-3.14	22,686.71	6.45	22.56	23.69	23.09	20.65	
		AMZN.O	亚马逊(AMAZON)	236.17	232.76	-1.44	24,667.27	3.87	41.63	51.61	39.22	35.96	
		META.O	脸书(META PLATFORMS)	704.87	719.80	2.12	18,237.28	11.09	29.25	23.27	23.79	27.92	
		NFLX.O	奈飞(NETFLIX)	1,011.11	1,008.08	-0.30	4,312.13	11.06	49.50	39.40	43.77	40.60	
		PDD.O	拼多多	110.14	117.48	6.66	1,631.53	3.07	10.56	22.94	0.00	8.60	
		NTES.O	网易	101.27	105.49	4.17	667.81	4.49	17.25	14.46	0.00	15.70	
		BIDU.O	百度	88.63	89.44	0.91	319.64	1.64	10.38	14.52	0.00	10.70	
		TCOM.O	携程网	68.82	65.69	-4.55	455.02	6.18	19.41	17.58	0.00	17.91	
		BABA.N	阿里巴巴	99.28	112.78	13.60	2,559.93	1.95	21.87	18.64	17.91	16.31	
		9988.HK	阿里巴巴-SW	97.55	104.90	7.53	19,938.99	1.87	20.95	18.53	17.79	15.58	
	0700.HK	腾讯控股	417.80	427.40	2.30	39,230.29	5.50	20.84	21.90	0.00	17.38		
	80700.HK	腾讯控股-R	389.80	400.00	2.62	36,121.29	5.71	21.63	21.86	0.00	17.67		
	9999.HK	网易-S	158.60	162.00	2.14	5,201.47	4.44	17.05	13.96	0.00	15.47		
	9888.HK	百度集团-SW	87.80	88.75	1.08	2,489.63	1.68	10.61	14.48	0.00	10.91		
89888.HK	百度集团-SWR	82.00	83.40	1.71	2,292.33	1.75	11.05	14.50	0.00	11.13			
1024.HK	快手-W	45.40	45.80	0.88	1,971.16	1.44	11.87	32.60	0.00	9.34			
81024.HK	快手-WR	42.45	42.90	1.06	1,814.94	1.50	12.33	32.58	0.00	9.50			
9626.HK	哔哩哔哩-W	139.90	137.70	-1.57	580.08	2.06	-19.14	-7.41	0.00	121.15			
2518.HK	汽车之家-S	54.55	53.00	-2.84	269.98	3.40	13.93	12.86	0.00	13.72			
9898.HK	微博-SW	77.10	81.50	5.71	198.30	1.45	6.80	7.52	0.00	6.73			
软件及应用	美股	MSFT.O	微软(MICROSOFT)	413.29	411.44	-0.45	30,586.38	11.68	32.98	38.62	35.56	32.20	
		SNOW.N	SNOWFLAKE	188.93	183.55	-2.85	605.90	17.75	-53.74	-82.25	-60.96	0.00	
		ORCL.N	甲骨文(ORACLE)	171.66	177.19	3.22	4,955.93	9.02	42.64	34.08	44.53	41.65	
		CRM.N	赛富时(SALESFORCE)	347.93	324.40	-6.76	3,104.51	8.35	52.31	1224.61	77.36	51.30	
		ADBE.O	奥多比(ADOBE)	437.63	458.82	4.84	1,997.24	9.29	35.92	50.01	35.27	28.80	
		INTU.O	财捷(INTUIT)	588.26	585.88	-0.40	1,640.67	9.89	56.21	73.39	59.40	47.02	
		SNPS.O	新思科技(SYNOPSYS)	528.42	529.14	0.14	817.94	13.35	36.14	63.63	33.15	44.31	
		CDNS.O	铿腾电子(CADENCE)	302.65	303.87	0.40	833.40	19.14	80.20	71.17	0.00	0.00	
		ADSK.O	欧特克(AUTODESK)	311.36	299.72	-3.74	644.40	10.81	59.06	63.29	70.14	62.57	
		U.N	Unity	21.11	19.50	-7.63	78.55	4.00	-9.89	-18.90	0.00	-16.54	
	0020.HK	商汤-W	1.65	1.71	3.64	632.83	15.55	-10.00	-5.46	0.00	-21.54		
	80020.HK	商汤-WR	1.53	1.61	5.23	582.67	16.04	-10.32	-5.46	0.00	-22.02		
	3888.HK	金山软件	42.70	43.30	1.41	578.06	5.33	40.20	61.69	0.00	28.01		
	0268.HK	金蝶国际	11.08	13.26	19.68	475.51	7.16	-300.95	-178.41	0.00	278.11		
9878.HK	汇通达网络	14.08	15.16	7.67	85.29	0.11	23.71	34.57	0.00	12.44			
3650.HK	KEEP	5.17	7.13	37.91	37.48	1.57	-13.54	6.00	0.00	1396.80			
0354.HK	中国软件国际	5.63	6.28	11.55	171.57	0.93	24.16	22.39	0.00	17.85			
1357.HK	美图公司	4.58	4.59	0.22	209.21	6.28	42.05	38.62	0.00	23.93			
3896.HK	金山云	8.20	8.90	8.54	338.67	4.20	-14.87	-3.18	0.00	-53.70			
2013.HK	微盟集团	1.93	2.35	21.76	84.97	4.11	-9.05	-9.62	0.00	-81.96			
1675.HK	亚信科技	6.01	6.94	15.47	64.91	0.78	22.98	13.69	0.00	9.21			
2121.HK	创新奇智	6.20	6.12	-1.29	34.58	2.25	-5.42	-8.15	0.00	-9.80			
2400.HK	心动公司	26.15	29.70	13.58	145.86	3.45	417.71	-53.48	0.00	16.58			
0777.HK	网龙	10.26	10.94	6.63	58.12	0.78	11.79	10.84	0.00	4.65			

资料来源：WIND, 中国银河证券研究院

二、行业要闻

（一）算力及终端

【CoreWeave 推出 NVIDIA GB200 NVL72 实例，推动 AI 推理模型发展】

CoreWeave 推出了基于 NVIDIA GB200 NVL72 的新实例，成为首家提供 NVIDIA Blackwell 平台的云服务提供商。这些实例配备了 72 个 NVIDIA Blackwell GPU 和 36 个 NVIDIA Grace CPU，并通过 NVIDIA Quantum-2 InfiniBand 网络连接，支持最大 110,000 个 GPU 的扩展，能够提供构建和部署下一代 AI 推理模型和智能体所需的规模与性能。NVIDIA GB200 NVL72 是一种液冷机架级解决方案，采用 72-GPU NVLink 域，使得 72 个 GPU 可以作为一个巨大的单一 GPU 运行，支持高效的推理任务。

NVIDIA Blackwell 平台带来了多项技术突破，如第五代 NVLink 提供每个 72-GPU NVLink 域高达 130TB/s 的 GPU 带宽，第二代 Transformer Engine 则支持 FP4，提升 AI 性能的同时保持高精度。CoreWeave 为 Blackwell 量身定制的云服务套件，包括 CoreWeave Kubernetes Service，优化了工作负载调度，并通过 Slurm on Kubernetes (SUNK) 实现智能工作负载分配。CoreWeave 的 Observability 平台还提供实时的 NVLink 性能、GPU 利用率和温度监控。

通过与 NVIDIA 的合作，CoreWeave 推出的这些实例使企业能够在云端部署最先进的 AI 推理模型和智能体，为下一代 AI 应用提供强大支持。客户可以通过 CoreWeave Kubernetes Service 在 US-WEST-01 区域启用 GB200 NVL72 实例，获取所需的规模和性能，推动 AI 领域的进一步发展。。

【“RTX 50 系列显卡加速生成 AI 本地处理”】

NVIDIA 的 GeForce RTX 5090 和 5080 显卡基于突破性的 Blackwell 架构，结合 NVIDIA DLSS 4 技术，提供高达 8 倍的帧率提升，同时借助 NVIDIA Reflex 2 技术降低延迟，并通过 NVIDIA RTX 神经着色器增强图形保真度。这些显卡专为加速最新的生成性 AI 工作负载而设计，提供高达 3352 万亿次 AI 操作 (TOPS)，为 AI 爱好者、游戏玩家、创作者和开发者带来前所未有的体验。

在 CES 展会上，NVIDIA 推出了 NIM 微服务和 AI 蓝图，旨在帮助开发者和爱好者利用 RTX 显卡加速 AI 开发。NIM 微服务是预打包的生成性 AI 模型，能够帮助开发者轻松启动 AI 项目并快速迭代，同时充分利用 RTX 显卡的强大性能。AI 蓝图则是展示如何使用 NIM 微服务构建下一代 AI 应用的参考项目，优化了 GeForce RTX 50 系列显卡的性能。这些技术可以无缝协作，帮助开发者在 AI PC 上快速构建和交付先进的 AI 体验。

NVIDIA 的 Tensor 核心在 Blackwell 架构中得到进一步强化，提供了 3352 TOPS 的 AI 计算能力，支持更复杂的 AI 任务并能够同时运行多个 AI 模型。通过采用 FP4 量化技术，AI 模型的大小减少了 60%，性能翻倍，能够在 GeForce RTX 5090 显卡上实现更快的 AI 处理，显著提高了内容创作的效率。NIM 微服务和 AI 蓝图结合 NVIDIA 的最新技术，使得生成性 AI 不再仅限于云端，而是可以在 RTX PC 上本地加速运行，带来更智能、更快速的 AI 体验。

（二）大模型及云应用

【DeepSeek：AI 新纪元的破局者】

DeepSeek AI 助手一经推出已成为全球增速最快的 AI 应用之一。上线 20 天内，其日活跃用户突破 2000 万，下载量在 18 天内达到 1600 万次，几乎是 ChatGPT 同期的两倍，成为 AI 领域的焦点。同时带来资本市场震动，多个科技巨头股价大幅下跌，英伟达的市值一夜蒸发近

6000 亿美元。

DeepSeek 的突破不仅体现在用户增长上，更在技术上打破了高端硬件的限制。其以不到 600 万美元的成本训练出与 ChatGPT-4o 相媲美的 AI 模型，训练效率是 Meta 的 11 倍。在推理层面，DeepSeek-R1 的推理成本仅为 OpenAI 的三十分之一。这一创新使得 AI 模型的开发成本大幅降低，促使更多企业和开发者参与 AI 创新，推动了行业的技术创新和成本重构。

此外，DeepSeek 的技术也在多个行业得到了广泛应用，特别是在电商、教育、智能硬件等领域，提升了运营效率和用户体验。通过低成本的 AI 技术，DeepSeek 为中小企业和开发者提供了参与 AI 开发的机会，降低了市场准入门槛，激发了创新活力。DeepSeek 的崛起不仅改变了 AI 行业的成本结构，也推动了 AI 应用的普及和技术的进步，未来有望在全球范围内持续引领 AI 领域的变革。

DeepSeek 火爆全球并影响全球人工智能产业链格局，以 AI Agent 为代表的 AI 应用及端侧 AI 有望迎来爆发期。2 月 9 日，沙特阿美旗下 Aramco Digital 宣布，DeepSeek 在其位于沙特阿拉伯达曼的数据中心开始运营。而此前 2 月 6 日，中国移动、中国联通、中国电信联合宣布已完成 DeepSeek-R1 及 V3 模型在云计算平台的全面接入。DeepSeek AI 助手一经推出已成为全球增速最快的 AI 应用之一，目前日活用户量已经突破 3000 万。

我们认为：DeepSeek 通过架构优化和训练方法革新降低单次训练和推理的算力需求，压缩成本至行业极低水平，但并未打破 Scaling Laws 的本质与创新边界，算力需求依然长期增长，由于低成本将大幅推动应用普及，未来推理算力将接棒训练算力成为算力消耗主力，全球 AI+行业应用、端侧 AI 预计将迎来新一轮繁荣期。

【OpenAI 发布 o3-mini】

OpenAI o3-mini 是一款高效、低成本的小型推理模型，在科学、数学和编程等 STEM 领域表现尤为出色。相较于前代 o1-mini，新模型不仅在推理能力上取得了显著提升，还具备更低的延迟和更高的速率限制。o3-mini 现已在 ChatGPT 和 API 中上线，支持函数调用、结构化输出以及开发者消息等功能，并允许用户根据需求选择低、中、高三种推理强度，以在复杂任务与响应速度之间取得平衡。

在性能方面，o3-mini 在数学、编程竞赛及博士级科学问题上的表现优于 o1-mini，甚至在高推理强度下超越了 o1。在 AIME 2024 竞赛数学题、GPQA 博士级科学测试及 Codeforces 竞赛编程任务中，o3-mini 均取得了更高的准确率和 Elo 评分。此外，o3-mini 在 LiveBench 代码生成、软件工程 (SWE-bench Verified) 等任务上的表现同样优异，能够提供更清晰、准确的回答，并显著减少重大错误。根据外部专家测试结果，相较于 o1-mini，o3-mini 的回答更受偏好，错误率降低了 39%。

o3-mini 现已向 ChatGPT Plus、Team 和 Pro 用户开放，并将很快面向企业用户推出。相较于 o1-mini，新模型在推理能力更强的同时，响应速度提升 24%，且 Plus 和 Team 账户的使用上限提高至每天 150 条。此外，o3-mini 现已支持搜索功能，可提供带有来源链接的最新信息。

【阿里通义千问全新视觉理解模型 Qwen2.5-VL】

阿里云通义团队发布全新的视觉模型 Qwen2.5-VL，推出了 3B、7B 和 72B 三个版本，其中旗舰版 Qwen2.5-VL-72B 在 13 项权威评测中获得视觉理解冠军，超越了 GPT-4o 和 Claude3.5。新模型支持更精准的图像解析和超长视频理解，能够无需微调便执行复杂操作，如手机和电脑的多步骤任务。Qwen2.5-VL 的视觉知识解析能力也实现了突破，能够识别图像中的布局结构、文本、图表等复杂内容，提升了 OCR 能力和文档解析能力，精准识别并抽取关键信息。

此次升级使得 Qwen2.5-VL 在多个任务中表现卓越，特别是在大学水平的问答、数学、文档理解、视觉问答及视频理解等领域，全面超越 GPT-4o 和 Claude3.5。通过引入动态帧率训练和时间编码技术，Qwen2.5-VL 在视频理解上取得了显著进步，可以准确处理小时级别的长视频内容，并高效提取关键信息。此外，Qwen2.5-VL 还具备作为视觉智能体的能力，可以在无需特定任务微调

的情况下，直接操控手机和电脑，执行如查询天气、订票等复杂操作，提供更智能化的自动化处理功能。

在模型技术上，Qwen2.5-VL 增强了对时间和空间尺度的感知，并优化了网络结构，提高了模型效率。通过创新的视觉编码器设计和结构优化，Qwen2.5-VL 进一步提升了其视觉编解码能力，为开发者提供了更高效、更灵活的多模态 AI 应用工具。

【Gemini 2.0 Flash 更新及新模型发布】

去年 12 月，谷歌发布了 Gemini 2.0 Flash 的实验版本，正式开启了代理型 AI 的新时代。Gemini 2.0 Flash 是谷歌为开发者群体打造的高效主力模型，具有低延迟、高性能等优势。今年早些时候，谷歌在 Google AI Studio 中更新了 2.0 Flash Thinking Experimental，通过将 Flash 模型的速度与复杂问题的推理能力相结合，进一步提高了性能表现。上周，谷歌面向桌面及移动设备端的全体 Gemini 应用用户发布了 2.0 Flash 更新版本，希望帮助更多人以全新方式使用 Gemini 进行创作、互动和协作。如今，谷歌将通过 Google AI Studio 和 Vertex AI 中的 Gemini API 向公众发布更新之后的 Gemini 2.0 Flash。开发人员现已可以使用 2.0 Flash 模型构建生产级应用程序。

此外，谷歌还发布了 Gemini 2.0 Pro 的实验版本，这是谷歌旗下迄今为止编码性能最强、最善于处理复杂提示词的大模型。除了在 Google AI Studio 和 Vertex AI 当中使用之外，Gemini 2.0 Pro 也将在 Gemini 应用中面向 Gemini Advanced 用户开放。同时，谷歌将在 Google AI Studio 和 Vertex AI 中公开预览迄今为止最具成本效益的模型方案 Gemini 2.0 Flash-Lite。最后，2.0 Flash Thinking Experimental 将被添加在桌面和移动设备端的模型下拉菜单中，以供 Gemini 应用用户随时使用。以上提到的所有发布模型都将支持带有文本输出的多模态输入，且在未来几个月的通用版本中还将支持更多模态。

【升级版 Claude 3.5 Sonnet 在 AWS 悉尼区域上线】

Anthropic 公司发布了最新升级的 Claude 3.5 Sonnet 模型，该模型现已在 AWS 亚太（悉尼）区域的 Amazon Bedrock 上全面上线。Claude 3.5 Sonnet 提供先进的编码能力和复杂问题处理能力，特别在编程和解决方案生成领域表现突出，同时在理解上下文和推理能力上也有显著提高。该模型还支持多模态输入，能够有效处理图像数据，尤其擅长解析图表和图像中的文本信息。

此外，Claude 3.5 Sonnet 还具备了一项实验性的新功能，能够通过计算机操作（如键盘敲击和鼠标点击）来完成需要多步骤的任务，例如在线研究和报销处理，这为开发人员提供了新的使用场景，如自动化研究和软件测试。AWS 与澳大利亚政府的合作进一步强化，Claude 3.5 Sonnet 的本地化部署为澳大利亚企业和公共部门提供了更强的 AI 支持，帮助它们满足数据本地化和合规要求，同时推动技术创新和经济发展。

这款升级的模型已被包括联邦银行、nib 集团、One New Zealand 和墨尔本大学在内的本地客户所青睐，能够帮助它们提升生成 AI 应用的测试、构建和部署效率。AWS 还宣布，将继续加大在澳大利亚的投资，并通过 Amazon Bedrock 提供更多先进的 AI 解决方案，以支持澳大利亚各行业的数字化转型。

三、风险提示

技术迭代不及预期风险；科技巨头竞争加剧风险；法律监管风险；供应链风险；下游需求不及预期风险。

分析师承诺及简介

本人承诺以勤勉的执业态度，独立、客观地出具本报告，本报告清晰准确地反映本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告的具体推荐或观点直接或间接相关。

吴砚靖 TMT/科创板研究负责人，北京大学软件项目管理硕士，10 年证券分析从业经验，历任中银国际证券首席分析师，国内大型知名 PE 机构研究部执行总经理。具备一二级市场经验，长期专注科技公司研究。

免责声明

本报告由中国银河证券股份有限公司（以下简称银河证券）向其客户提供。银河证券无需因接收人收到本报告而视其为客户。若您并非银河证券客户中的专业投资者，为保证服务质量、控制投资风险、应首先联系银河证券机构销售部门或客户经理，完成投资者适当性匹配，并充分了解该项服务的性质、特点、使用的注意事项以及若不当使用可能带来的风险或损失。

本报告所载的全部内容只提供给客户做参考之用，并不构成对客户投资咨询建议，并非作为买卖、认购证券或其它金融工具的邀请或保证。客户不应单纯依靠本报告而取代自我独立判断。银河证券认为本报告资料来源是可靠的，所载内容及观点客观公正，但不担保其准确性或完整性。本报告所载内容反映的是银河证券在最初发表本报告日期当日的判断，银河证券可发出其它与本报告所载内容不一致或有不同结论的报告，但银河证券没有义务和责任去及时更新本报告涉及的内容并通知客户。银河证券不对因客户使用本报告而导致的损失负任何责任。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的银河证券网站以外的地址或超级链接，银河证券不对其内容负责。链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

银河证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。银河证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

银河证券已具备中国证监会批复的证券投资咨询业务资格。除非另有说明，所有本报告的版权属于银河证券。未经银河证券书面授权许可，任何机构或个人不得以任何形式转发、转载、翻版或传播本报告。特提醒公众投资者慎重使用未经授权刊载或者转发的本公司证券研究报告。

本报告版权归银河证券所有并保留最终解释权。

评级标准

评级标准	评级	说明
评级标准为报告发布日后的6到12个月行业指数（或公司股价）相对市场表现，其中：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，北交所市场以北证50指数为基准，香港市场以恒生指数为基准。	行业评级	推荐：相对基准指数涨幅10%以上
		中性：相对基准指数涨幅在-5%~10%之间
		回避：相对基准指数跌幅5%以上
	公司评级	推荐：相对基准指数涨幅20%以上
		谨慎推荐：相对基准指数涨幅在5%~20%之间
		中性：相对基准指数涨幅在-5%~5%之间
	回避：相对基准指数跌幅5%以上	

联系

中国银河证券股份有限公司 研究院

深圳市福田区金田路3088号中洲大厦20层

上海浦东新区富城路99号震旦大厦31层

北京市丰台区西营街8号院1号楼青海金融大厦

公司网址：www.chinastock.com.cn

机构请致电：

深广地区：程曦 0755-83471683 chengxi_yj@chinastock.com.cn

苏一耘 0755-83479312 suyiyun_yj@chinastock.com.cn

上海地区：陆韵如 021-60387901 luyunru_yj@chinastock.com.cn

李洋洋 021-20252671 liyangyang_yj@chinastock.com.cn

北京地区：田薇 010-80927721 tianwei@chinastock.com.cn

褚颖 010-80927755 chuying_yj@chinastock.com.cn