仅供内部参考，请勿外传！

行业评级	推荐（维持）
报告日期	2026 年 09 月 13 日

相关研究

【兴证计算机】超节点：大集群算力之基，国产互联启新篇-2026.09.08

【兴证计算机】多模态：商业化高歌猛进，模型能力持续提升（8.31-9.6）-2026.09.06

【兴证计算机】具身智能：ICL 能力达临界，数采潜力巨大（8.24-8.30）-2026.08.30

分析师：蒋佳霖
S0190515050002
jiangjialin@xyzq.com.cn

分析师：杨本鸿
S0190522080001
yangbenhong@xyzq.com.cn

AGI：SOTA 模型泛化突破，产业曙光初现

投资要点：

- 本周申万计算机指数下跌 4.61%，跑输万得全 A 指数 3.55 个百分点，位列 30/31；交易额 0.43 万亿元，位列 8/31。
- **核心观点**：大模型迈向 AGI，提振算力及应用需求
- **DS V4.1 Flash 登场，大模型资本化持续加速。**本周，DeepSeek 正式发布 V4.1 Flash，在性能、定价、速度等均优于 V4 Pro，并将 Pro 版路由切换至 Flash；同时，公司正加快资本化步伐，有望启动 IPO 进程。在海外，据路透社，Anthropic 计划 IPO 最高募资 1000 亿美元、估值或达 2 万亿美元，英伟达拟作为基石投资 100 亿美元。
- **大模型进阶拉动算力需求，CPU 涨价趋势延续。**本周，OpenAI 宣布因 GPT-6 Astra 需求快速增长，暂停接受每月 200 美元的 Pro 20X 新增订阅，高阶模型上线形成对算力需求的显著拉动。同时，英特尔计划于 10 月上旬再次上调电脑 CPU 价格，涨幅预计达 10%，进一步体现 Agent 带来的 AI 基础设施需求。在国内，除 CSP 外的政府开支亦有望提速。下周，随着华为全联接大会的召开，持续关注国产算力链的产业动态。
- **继续关注"AI+行业"的政策红利与受益龙头。**近期，AI 软件政策密集出台：9 月初，工信部发布“AI+软件”专项方案，推动软件产业智能化升级；"AI+消费"、"数字化绿色化"等政策驱动 AI 赋能消费、制造业等需求扩张。同时，GPT-6 带来 AGI 热度上行，关注 Agentic 能力边界拓展，重视办公、工业、设计等企业级 AI 领域的潜在需求。
- **本文第二章，从 AGI 视角对大模型产业近况进行了重点跟踪。**
- **建议持续关注：**
- 1）芯片层：海光信息、寒武纪、摩尔线程、沐曦股份、壁仞科技、天数智芯、盛科通信、概伦电子、中国长城、英伟达、AMD、英特尔、ARM、博通；
- 2）AI 模型：中控技术、智谱、Minimax、科大讯飞、索辰科技、群核科技、五一视界、Momenta；
- 3）AI 应用：金山办公、迈富时、泛微网络、合合信息、汉得信息、税友股份、新大陆、卓易信息、福昕软件、鼎捷数智、金蝶国际、能科科技、虹软科技、德赛西威、海康威视、大华股份、萤石网络、中科创达、星环科技、汇量科技、同花顺、恒生电子、顶点软件、南网数字、科远智慧、剂泰科技、金斯瑞生物科技、晶泰科技、英矽智能；
- 4）AI Infra：浪潮信息、紫光股份、中科曙光、达梦数据、宝信软件、工业富联、澜起科技、神州数码、金山云、优刻得、网宿科技、数据港、首都在线。

风险提示：板块业绩存在不达预期的风险；下游需求存在景气度不稳定的风险；科技创新可能带来短期业绩下降。

仅供内部参考，请勿外传！

目录

1、行业周观点.....3
2、AGI：SOTA 模型泛化突破，产业曙光初现.....4
2.1 前沿大模型密集迭代，AGI 有望加速到来.....4
2.2 从 AGI 到 ASI，AI 能力仍有巨大提升潜力.....6
2.3 投资建议与相关标的.....8
3、风险提示.....8

图目录

图 1、 GPT-6 在 ECI 指数中超越其他前沿模型.....4
图 2、 GPT-6 模型能力幻觉率显著降低.....4
图 3、 Fable5.1 准确率提升，成本降低.....5
图 4、 Fable 5.1 在各项基准测试中超越前代.....5
图 5、 Muse Spark 1.3 模型编码能力大幅提升.....5
图 6、 DeepSeek V4.1 Flash 缓存占用显著下降.....5
图 7、 AI 的发展可划分为三阶段.....6
图 8、 AGI 当前仍处于 L2 快速发展阶段.....7

表目录

表 1、 向 AGI/ASI 发展潜在的瓶颈与应对方案.....7

仅供内部参考，请勿外传！

1、行业周观点

核心观点：大模型迈向 AGI，提振算力及应用需求

DS V4.1 Flash 登场，大模型资本化持续加速。本周，DeepSeek 正式发布 V4.1 Flash，在性能、定价、速度等均优于 V4 Pro，并将 Pro 版路由切换至 Flash；同时，公司正加快资本化步伐，有望启动 IPO 进程。在海外，据路透社，Anthropic 计划 IPO 最高募资 1000 亿美元、估值或达 2 万亿美元，英伟达拟作为基石投资 100 亿美元。

大模型进阶拉动算力需求，CPU 涨价趋势延续。本周，OpenAI 宣布因 GPT-6 Astra 需求快速增长，暂停接受每月 200 美元的 Pro 20X 新增订阅，高阶模型上线形成对算力需求的显著拉动。同时，英特尔计划于 10 月上旬再次上调电脑 CPU 价格，涨幅预计达 10%，进一步体现 Agent 带来的 AI 基础设施需求。在国内，除 CSP 外的政府开支亦有望提速。下周，随着华为全联接大会的召开，持续关注国产算力链的产业动态。

继续关注"AI+行业"的政策红利与受益龙头。近期，AI 软件政策密集出台：9 月初，工信部发布“AI+软件”专项方案，推动软件产业智能化升级；"AI+消费"、"数字化绿色化"等政策驱动 AI 赋能消费、制造业等需求扩张。同时，GPT-6 带来 AGI 热度上行，关注 Agentic 能力边界拓展，重视办公、工业、设计等企业级 AI 领域的潜在需求。

建议持续关注：

1）芯片层：海光信息、寒武纪、摩尔线程、沐曦股份、壁仞科技、天数智芯、盛科通信、概伦电子、中国长城、英伟达、AMD、英特尔、ARM、博通；

2）AI 模型：中控技术、智谱、Minimax、科大讯飞、索辰科技、群核科技、五一视界、Momenta；

3）AI 应用：金山办公、迈富时、泛微网络、合合信息、汉得信息、税友股份、新大陆、卓易信息、福昕软件、鼎捷数智、金蝶国际、能科科技、虹软科技、德赛西威、海康威视、大华股份、萤石网络、中科创达、星环科技、汇量科技、同花顺、恒生电子、顶点软件、南网数字、科远智慧、剂泰科技、金斯瑞生物科技、晶泰科技、英矽智能；

4）AI Infra：浪潮信息、紫光股份、中科曙光、达梦数据、宝信软件、工业富联、澜起科技、神州数码、金山云、优刻得、网宿科技、数据港、首都在线。

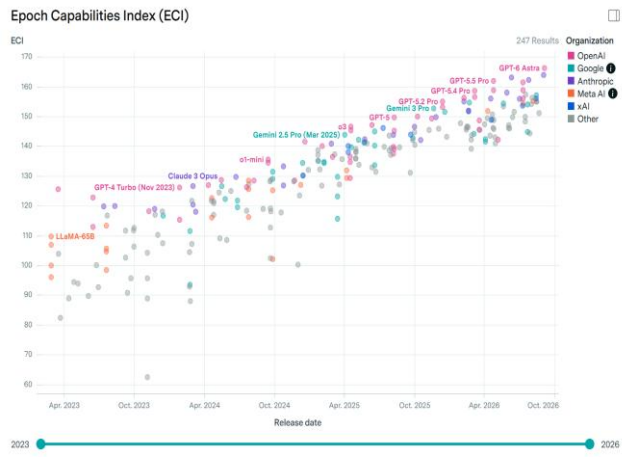
仅供内部参考，请勿外传！

2、AGI：SOTA 模型泛化突破，产业曙光初现

2.1 前沿大模型密集迭代，AGI 有望加速到来

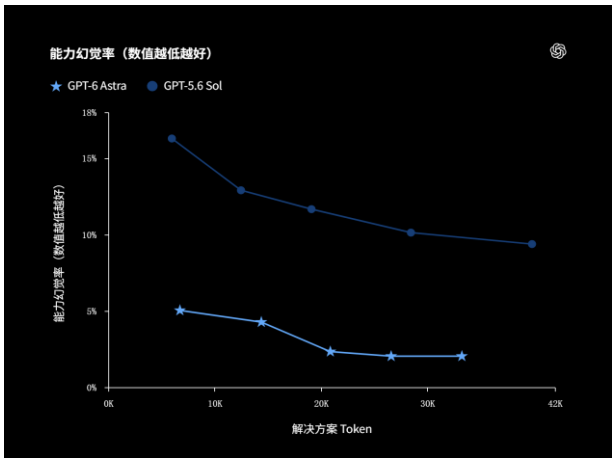
以 GPT-6 Astra 为代表的 SOTA 模型正快速迈向 AGI。目前，AGI 正成为大模型产业新的聚焦点。OpenAI 将 GPT-6 Astra 定义为"新一代智能"，Astra 在计算机操作、网页浏览、软件工程、网络安全、科学研究等领域均达 SOTA，更将智能延伸至物理世界，具备操控电脑、驱动机械臂、生成可 3D 打印模型、完成 PCB 电路板布局等能力，在各项基准测试均保持领先。此外，Astra 是 Open AI 首个达"Critical"级网络安全能力的模型，并在素数间隔等数学难题上取得突破。模型支持 105 万 token 上下文，API 定价$10/$50 每百万 token，正逐步向全量用户开放。

图1、GPT-6 在 ECI 指数中超越其他前沿模型



数据来源：Epoch AI，兴业证券经济与金融研究院整理

图2、GPT-6 模型能力幻觉率显著降低

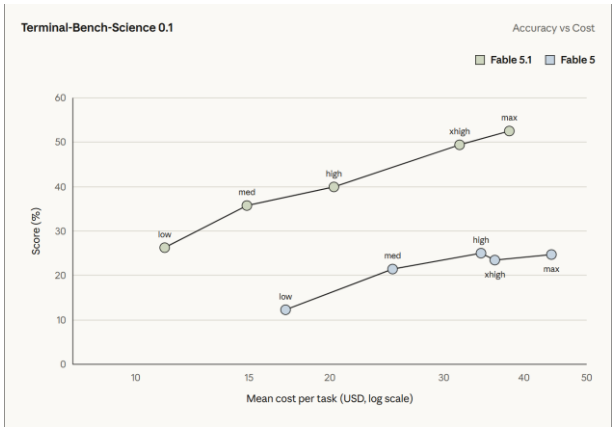


数据来源：Open AI，兴业证券经济与金融研究院整理

Anthropic 发布 Fable 5.1 和 Mythos 5.1，同底座领跑编程与科研。2026 年 9 月 1 日，Anthropic 发布 Claude Fable 5.1 和 Claude Mythos 5.1，两款模型共享同一底层模型，区别在于安全层级：Fable 5.1 面向公众广泛开放，Mythos 5.1 则仅向经过审核的网络安全和生命科学机构开放。官方 System Card 指出，该模型在编程、知识工作和问题解决领域推进了前沿，并在新颖数学和科学推理能力上有所提升。在 Terminal-Bench Science 基准上，Fable 5.1 得分 52.6%，较前代 Fable 5 的 24.7%实现翻倍。同步推出的还有反蒸馏安全机制——Messages API 将思考块与产生它的对话上下文绑定，以防范未经授权的大规模模型蒸馏。

仅供内部参考，请勿外传！

图3、Fable5.1 准确率提升，成本降低



数据来源：Anthropic，兴业证券经济与金融研究院整理

图4、Fable 5.1 在各项基准测试中超越前代

	Fable 5.1	Fable 5	Opus 5	GPT-5.6 Sol
Agentic scientific research Terminal-Bench-Science 0.1 [1]	**52.6%**	24.7%	29.0%	22.4%
Agentic coding Terminal-Bench 4.0	**55.8%** 60.9% (Mythos 5.1)	42.0%	52.3%	37.3%
Knowledge work GDPval-AA v2	**1853**	1723	1824	1711
Computer use OSWorld 2.0 [2]	**77.9%** partial	72.9% partial	75.4% partial	— partial
Computer use OSWorld 2.0	**41.7%** strict	36.1% strict	39.6% strict	— strict
Multidisciplinary reasoning Humanity's Last Exam	**60.9%** no tools	57.8% no tools	56.6% no tools	— no tools
	65.0% with tools	63.8% with tools	63.6% with tools	— with tools
Business workflows AutomationBench	**31.4%**	17.1%	26.9%	19.6%
Agentic coding CursorBench 3.2.0	**73.4%**	70.5%	70.0%	67.2%

数据来源：Anthropic，兴业证券经济与金融研究院整理

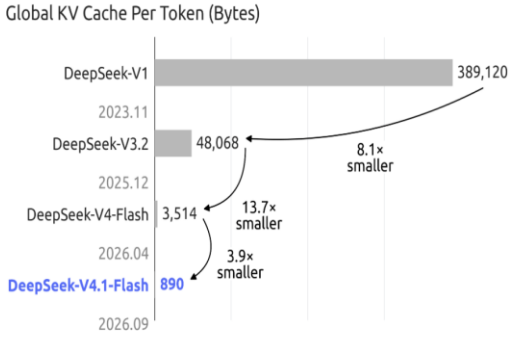
Meta、谷歌、DeepSeek 新模型同步密集发布，智能水平均有较大提升。1）9 月 2 日，Meta 发布 Muse Spark 1.3，五个月内第四次迭代，主打长周期智能体与编码能力，DeepSWE 1.1 达 75.4%、Terminal-Bench 2.1 达 88.8%，工具调用减少 20%、token 消耗减少 25%，在复杂多步骤任务上的自主执行与自我纠错能力显著增强，标志着个人超级智能加速落地；**2**）9 月 2 日，谷歌推出 Gemini 3.8 Flash 及网络安全特化版 3.8 Flash Cyber，六周内第三次迭代，以 Flash 级延迟和成本在 DeepSWE 长周期软件工程基准上超越多数更大的前沿模型，Cyber 版则在漏洞检测与自动修复上达到前沿水平，展现专业化垂直场景渗透能力；**3**） 9 月 10 日，DeepSeek 发布 V4.1 Flash，采用非对称 CED 架构，以小尺寸模型在性能、速度、成本上全面超越 DeepSeek V4 Pro，GPQA Diamond 达 90+，并以 MIT 协议开源，证明 AGI 能力突破不再单纯依赖参数堆叠，架构创新有望成为智能水平跃升的新方向。**4**）9 月 11 日，月之暗面 Kimi K2.8 Preview 上线 KimiCode，模型 ID 不变无需修改配置。其综合性能接近 K3，编码、智能体及思考效率提升，全会员开放 1M 上下文，支持三档思考强度调节。

图5、Muse Spark 1.3 模型编码能力大幅提升

	Benchmark	Muse Spark 1.3 (max)	Muse Spark 1.2 (xhigh)	GPT 5.6 Sol (max)	Opus 5 (max)
Agent	GDPVal-AA v2	1754	1615	1710	1824
	JobBench	64.9	61.6	45.4	65.7
	OSWorld 2.0	66.9	47.6	62.7	68.3
	DeepSearchQA	89.4	85.9	93.0	90.4
	Agentic IF Index (internal)	57.8	46.2	60.5	59.1
	AutomationBench	49.4	38.2	46.7	50.3
Long Context	MRCR 256K-512K	98.5	66.3	91.5	-
	MRCR 512K-1M	98.1	55.5	73.8	-
Coding	DeepSWE v1.1	75.4	55.0	73.0	74.0
	SWEAtlas CodeBase QnA	59.4	46.2	53.5	52.7
	Terminal-Bench 2.1	88.8	82.9	88.8	86.7

数据来源：IT 之家，兴业证券经济与金融研究院整理

图6、DeepSeek V4.1 Flash 缓存占用显著下降



数据来源：DeepSeek，兴业证券经济与金融研究院整理

仅供内部参考，请勿外传！

2.2 从 AGI 到 ASI，AI 能力仍有巨大提升潜力

ANI 到 AGI 强调智能“可迁移性”提升，AGI 到 ASI 强调智能“量级”跃迁。当前可将 AI 发展分为三个阶段，即 ANI（专用人工智能）、AGI（通用人工智能）、ASI（超级人工智能）。对比来看，ANI 和 AGI 的区别在于 AI 的智能水平，当前的主流 AI 模型和产品仍为 ANI，其在某个特定任务中表现出色，但迁移能力较差；AGI 则可以实现跨领域理解、学习、推理，并把在一个场景里学到的能力，迁移到另一个从没见过的场景里去解决问题，AGI 的智能水平与平均人类智力类似，强调解决任何陌生问题的通用认知能力，而非局限于通用目的（General-Purpose，即能解决多种问题），真正的 AGI 应该具备整套认知能力，包括：通用推理、通用学习、知识迁移、长远规划、自主解决问题、记忆、适应能力等。此外，广义 AGI 还应与物理世界有所交互，根据阿里巴巴吴泳铭，AGI 应能实现 Agent 完成数字世界工作，并通过接口连接真实世界，操作部分物理设备；AGI 进一步发展后可实现 ASI，其具备在绝大多数认知任务上显著超越最优秀的人类个体，甚至人类整个群体的智力水平。

图7、AI 的发展可划分为三阶段

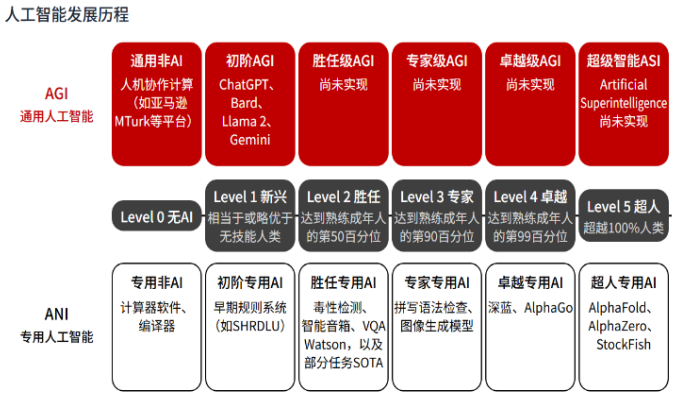


数据来源：AITNT，兴业证券经济与金融研究院整理

AGI 正处于蓬勃发展阶段，智能水平已可实现超越部分人类。根据 Google DeepMind《Position: Levels of AGI for Operationalizing Progress on the Path to AGI》，AI 的智能水平可分为 L0-L5 六个层级，从 L0 的无智能水平，到 L5 的超越全部人类。当前 ANI 已经实现 L5 级别的智能（AlphaFold 在精准预测蛋白结构的能力上已超越人类），而 AGI 则仍处于 L2 级别，即新兴发展阶段。以 ChatGPT 等产品为代表，其智能水平相当于或优于未受过特殊技巧训练的人类。若要达到 L3，即完全智能阶段，则智能水平至少应相当于受训人类的一半水平。此外，在此划分体系下，ASI 也可理解为 L5 级别，即智能水平超越全部人类的 AGI。

仅供内部参考，请勿外传！

图8、AGI 当前仍处于 L2 快速发展阶段



数据来源：Google DeepMind《Position: Levels of AGI for Operationalizing Progress on the Path to AGI》，2024，兴业证券经济与金融研究院整理

AGI 向下一阶段迈进仍需时日。据 Google DeepMind《From AGI to ASI》，当前模型智能虽显著提升，但发展仍有诸多瓶颈：算法与高质量数据供给存在局限，神经网络与随机梯度下降框架尚有短板；AI 难以挖掘训练数据里人类未知的抽象认知。同时物理端约束加剧，算力、供应链、能源、稀土及政策资金等，持续制约头部企业 AI 发展。长期看，AI 场景落地抬升 ROI，叠加算力、存储、能源效率改善，瓶颈有望逐步缓解。

表1、向 AGI/ASI 发展潜在的瓶颈与应对方案

潜在瓶颈方向	瓶颈具体描述	可能的应对方案
数据不足	模型在预训练、后训练、微调以及测试时自适应阶段，难以获得足量的高质量数据	合成数据、高保真仿真、自生成数据；能够提升数据利用效率的革新
经济与自然资源限制	若延续当前主流技术范式持续扩规模，所需的经济投入、技术要素（芯片、供应链）与自然资源（能源、适配的数据中心选址、稀土等）将持续增长	通过 AI 落地部署提升经济回报；依托 AI 研究提升算力、能源、数据的利用效率；开展大规模基础设施建设
神经网络范式局限性	仅依靠大规模预训练神经网络（叠加后训练、测试时缩放、提示脚手架、工具调用）与随机梯度下降，无法实现通用人工智能 AGI	持续开展 AI 研究，推动范式迭代与范式转换；即便能力不及 AGI 的 AI 系统，也可加速相关领域的研发工作
研发难度持续抬升	随着 AI 领域逐步成熟、持续推进 AI 研发所需投入或将显著增加。除显性资源外也包括更抽象的研发成本，例如在假设空间中开展更高强度的搜索	能力更强的 AI 系统能提升研发效率（减少所需实验的数量、提出更优假设等），同时提升数字科研主体的资源利用效率
抽象概念壁垒	当前 AI 系统主要基于人类已形成的抽象概念训练，这意味着 AI 难以从原始数据中自主提炼全新概念与抽象认知	持续的规模扩张（更多、更快算力）与多智能体集群模式；或需通过范式转换（交互式学习、强化学习）直接攻克这一壁垒
人为有意减速	恶意主体滥用、事故或重大风险、军事层面滥用、社会文化损害以及公众抵制等	经济与政治压力、国际竞赛格局

数据来源：Google DeepMind《From AGI to ASI》，2026，兴业证券经济与金融研究院整理



仅供内部参考，请勿外传！

2.3 投资建议与相关标的

近期海内外厂商模型密集发布，对 AGI 的探索快速推进，更多 AI 可用场景逐步打开。建议持续关注：

1）**模型端**：中控技术、海康威视、大华股份、MiniMax、智谱、阿里巴巴、腾讯、科大讯飞、索辰科技、万兴科技、群核科技、五一视界、剂泰科技等；

2）**算力端**：海光信息、寒武纪、浪潮信息、紫光股份、中科曙光、摩尔线程、沐曦股份、壁仞科技、天数智芯、盛科通信、工业富联、澜起科技、神州数码。

3、风险提示

板块业绩存在不达预期的风险；下游需求存在景气度不稳定的风险；科技创新可能带来短期业绩下降。

仅供内部参考，请勿外传！

分析师声明

本人具有中国证券业协会授予的证券投资咨询执业资格并登记为证券分析师，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收到任何形式的补偿。

投资评级说明

投资建议的评级标准	类别	评级	说明
报告中投资建议所涉及的评级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的 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。其中：沪深两市以沪深 300 指数为基准；北交所市场以北证 50 指数为基准；新三板市场以三板成指为基准；香港市场以恒生指数为基准；美国市场以标普 500 或纳斯达克综合指数为基准。	股票评级	买入	相对同期相关证券市场代表性指数涨幅大于 15%
		增持	相对同期相关证券市场代表性指数涨幅在 5%～15%之间
		中性	相对同期相关证券市场代表性指数涨幅在-5%～5%之间
		减持	相对同期相关证券市场代表性指数涨幅小于-5%
		无评级	由于我们无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使我们无法给出明确的投资评级
	行业评级	推荐	相对表现优于同期相关证券市场代表性指数
		中性	相对表现与同期相关证券市场代表性指数持平
		回避	相对表现弱于同期相关证券市场代表性指数

信息披露

本公司在知晓的范围内履行信息披露义务。客户可登录 www.xyzq.com.cn 内幕交易防控栏内查询静默期安排和关联公司持股情况。

本公司为澜起科技(688008)、金山办公(688111)、中控技术(688777)、寒武纪(688256)、海光信息(688041)、达梦数据(688692)做市商。但上述持仓不曾、不会、不将对研究业务的独立性、客观性产生影响。

使用本研究报告的风险提示以及法律声明

兴业证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

，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、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约，投资者自主作出投资决策并自行承担投资风险，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，任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以本公司向客户发布的本报告完整版本为准。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。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本报告所载资料的来源被认为是可靠的，但本公司不保证其准确性或完整性，也不保证所包含的信息和建议不会发生任何变更。本公司并不对使用本报告所包含的材料产生的任何直接或间接损失或与此相关的其他任何损失承担任何责任。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌，过往表现不应作为日后的表现依据；在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告；本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现。过往的业绩表现亦不应作为日后回报的预示。我们不承诺也不保证，任何所预示的回报会得以实现。分析中所做的回报预测可能是基于相应的假设。任何假设的变化可能会显著地影响所预测的回报。

本公司的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司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的意见或建议不一致的投资决策。

本报告并非针对或意图发送予或为任何就发送、发布、可得到或使用此报告而使兴业证券股份有限公司及其关联子公司等违反当地的法律或法规或可致使兴业证券股份有限公司受制于相关法律或法规的任何地区、国家或其他管辖区域的公民或居民，包括但不限于美国及美国公民（1934 年美国《证券交易所》第 15a-6 条例定义为本「主要美国机构投资者」除外）。

本报告的版权归本公司所有。本公司对本报告保留一切权利。除非另有书面显示，否则本报告中的所有材料的版权均属本公司。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。未经授权的转载，本公司不承担任何转载责任。

特别声明

在法律许可的情况下，兴业证券股份有限公司可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到兴业证券股份有限公司及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。

兴业证券研究

上海	北京	深圳
地址：上海浦东新区长柳路 36 号兴业证券大厦 15 层	地址：北京市朝阳区建国门外大街甲 6 号世界财富大厦 32 层 01-08 单元	地址：深圳市福田区皇岗路 5001 号深业上城 T2 座 52 楼
邮编：200135	邮编：100020	邮编：518035
邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn	邮箱：research@xyzq.com.cn