2026 年 11 月 11 日

证券研究报告•行业研究•传媒

AI 系列之模型速评：Google

Gemini 4 Carbon 面向付费用户开放，多工作流基准登顶并上线多智能体编排

投资要点

- **事件：**北京时间 11 月 11 日，Google 宣布向付费用户全面开放新一代旗舰推理模型 Gemini 4 Carbon（gemini-4-carbon），并在代理开发平台 Google Antigravity 同步上线多智能体协同功能 /swarm（预览版）。同日，Carbon 正式上线 GitHub Copilot 及 Cursor，Google 亦联合 Artificial Analysis 发布多模态知识智能体基准 MKA。

- **距先导内测仅两周，Carbon 定位“高难度长程任务与网安防御”，同步推出 /swarm 多智能体编排。**Carbon 基于约六周前发布的 Gemini 4 Argon 构建，曾于 10 月 28 日面向受邀网安人员小范围发布，本次全面开放仅隔两周。官方建议日常编程与知识工作优先使用高效率的 Argon，将 Carbon 定位于复杂多步编程、跨系统调试、科学探索及网安防御。据官方模型卡，Carbon 支持 100 万 tokens 上下文与最高 100 万 tokens 连续输出，支持文本、图像、音视频输入及三档可调推理深度（low/medium/high）。为发挥其长程能力，Antigravity 新增 /swarm 预览功能，可将单条提示词转化为由协调者、独立 Git 工作树并行执行者及对抗验证者组成的智能体团队，支持将 Carbon 绑定于顶层架构与审计角色，搭配轻量模型并行改代码。渠道方面，API 付费层（AI Studio、Android Studio，免费层不开放）、Gemini for business 及 GitHub Copilot（Pro+/Max/Business/Enterprise）已直接开放；C 端面向 Google AI Ultra 订阅用户在 Gemini App、搜索 AI Mode 及 Antigravity 推送；网安专用版 Carbon for cyber 继续通过 Fairwind 审核开放。

- **能力层面，Carbon 在工作流、长上下文及网安评测中领先，第三方综合榜单位列全球第二。**据 Google 与 CursorBench、Sugra 基准数据，Carbon 在智能体编程 CursorBench 4.0（60.4%）和终端工程 Terminal-Bench 4.0（65.2%）上均居第一，高于 Claude Fable 5.2（60.0% / 61.6%）与 Opus 5.5（57.8% / 59.6%）；业务流自动化 AutomationBench（53.9%）、搜索智能体 WANDR 及模拟经营 Vending-Bench 2（$16,484）同样位列榜首；256k－1M 长上下文 GraphWalks（96.2%）、工程图纸理解 Blueprint-Bench 2（58.6%）及数学推理 RiemannBench（80.0%）大幅领先竞品；网安三大基准 SEC-bench Pro（88.2%）、ExploitGym（54.8%）和 CWE-bench（71.0%）均夺冠。但并非全面占优：FrontierSWE v2（61.2%）、FrontierCode 2.0（40.6%）及桌面操作 OSWorld-2.1（81.5%）仍低于 Fable 5.2（分别为 69.9%、49.1%、89.3%）。第三方方面，据 Artificial Analysis 最新 v5.0.1 榜单，Carbon（High 档位）以 60 分位居综合智能指数第二，仅次于 Fable 5.2 Max（65 分），领先 Opus 5.5（59 分）、GPT-6 Astra（57 分）与 Argon（50 分）。

- **官方同步披露多项跨学科标志性案例，指向多智能体工程与科学发现能力。**据 Google 披露，Carbon 多智能体团队仅凭 7 份市政规划 PDF（206 张工程图）和单条提示词，经 10 轮夜间循环构建出 90.6 万三角形的办公楼交互式 3D 模型；结合 AlphaProof Code 框架在数周内生成数百万行 Lean 代码，完成对 Signal 加密协议核心代码的数学形式化证明（传统需数年）；在

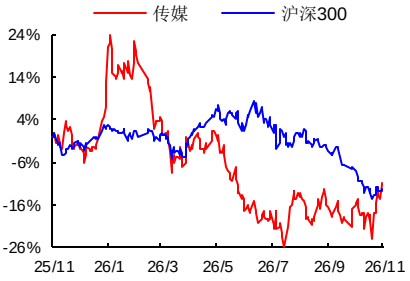
西南证券研究院

分析师：苟宇睿
执业证号：S1250524030001
电话：15882659190
邮箱：gyr@swsc.com.cn

分析师：潘妍洁
执业证号：S1250525100002
电话：023-65796461
邮箱：panyj@swsc.com.cn

分析师：刘言
执业证号：S1250515070002
电话：023-67791663
邮箱：liuyan@swsc.com.cn

行业相对指数表现



数据来源：聚源数据

基础数据

股票家数	129
行业总市值(亿元)	16,031.13
A 股市值(亿元)	15,925.56
行业市盈率 TTM	40.4
沪深 300 市盈率 TTM	13.4

相关研究

1. AI 系列之模型速评：Anthropic 启动 IPO 路演，Fable 5.2 登顶第三方智能榜单 (2026-11-11)
2. 传媒行业 2026 年 Q4 投资策略：Agent 迭代成熟，看好 AI 影视和 AI 营销 (2026-10-13)
3. 传媒行业 2026 年投资策略：AI 应用全面开花，游戏电影持续复苏 (2026-01-26)

CASP17 盲测集 7 个低同源蛋白家族中检索到的远缘同源物数量平均提升 40 倍；云安全平台 Wiz 利用其在 2 小时内挖出两家大型 SaaS 服务商的 0-day 漏洞。

- **首发期延续 $2/$10 折半促销价，标准单价与 Argon 持平（$4/$20），单任务成本约为 Fable 5.2 Max 的 29%。**据官方定价披露，Carbon 在 11 月 30 日前享首发半价，每百万 tokens 输入/输出为 $2/$10；12 月 1 日起恢复标准定价 $4/$20，与 Argon 及 Opus 5.5 持平，仅为 Fable 5.2 标准单价（$8/$40）的一半。据 Artificial Analysis，Carbon（High 档位）输出速度约 78 tokens/s，单任务平均输出 12.3 万 tokens（较 Opus 5.5 少 29%），标准单价下单任务平均成本仅 $7.76（Medium 档位 $6.59），仅为 Fable 5.2 Max（$26.54）的 29%。但在相同单价下，因 Carbon 思考链更长，实际账单仍高于 Argon。

- **第三方机构确认帕累托成本优势，企业客户反馈聚焦于交付物质量与零回归修复。**Artificial Analysis 指出，Carbon 处于智能指数与成本的帕累托前沿，显著补齐了 Google 在终端编程（Coding Agent 指数 59 分，仅次于 Opus 5.5 的 62 分）和演示文稿生成上的短板；在 MKA 评测中，Carbon（High）得分 54%，仅次于 Fable 5.2（Max）的 57%。据谷歌披露，Carbon 平均单任务成本约为榜首模型的六分之一。客户反馈方面：金融投研平台 Hebbia 称其构建的 Excel 模型与 PPT 为已测模型中最强；对冲基金 Balyasny 称其在内部 23 个模型配置中排名第一；OpenCode 工程师称其一夜间从零写出新 Rust UI 库并重写桌面应用；HubX 实测 1.4 万个测试代码库中的 16 个 Bug 修复实现 94% 解决率且零回归；SentinelOne 称其在内核驱动评估中额外发现 4 个经验证漏洞。

- **我们认为：**1）头部厂商旗舰模型迭代已压缩至“双周级”，Google（11 月 10 日开放 Carbon）与 Anthropic（11 月 9 日发布 Fable 5.2）正面交锋，且双方均采用“带护栏公众付费版 + 受审核机构专用网安/科研版”的双轨分发机制，验证了高风险前沿能力的商业化范式；2）Google 采取“与 Argon 同单价、按推理深度自然分层”的定价策略，Carbon High 档位在综合智能接近 Fable 5.2 Max（60 vs. 65）的同时，标准单任务成本（$7.76 vs. $26.54）低逾 70%，促销期成本优势更达近 7 倍，凸显自研 TPU 算力底座的推理经济性壁垒；3）Antigravity /swarm 将“重型模型负责顶层架构审计 + 轻量模型并行编码”的异构多智能体协同产品化，10 轮夜间循环或百万行形式化证明使单任务 tokens 消耗较常规对话跃升百倍以上，持续拉动推理算力与高阶订阅变现；4）模型在金融建模、工程图纸理解与 0-day 漏洞挖掘上的落地，表明竞争正从单点跑分转向端到端工程构件交付。

- **风险提示：**大模型迭代不及预期风险；12 月促销期结束后调用成本翻倍导致需求波动风险；行业竞争加剧风险；海外监管政策变化风险；第三方评测口径变化导致数据不可比风险。

图 1：Google 官方评测：2026 年 10 月 Gemini 4 Carbon 多维度基准测试结果对比

	Benchmark		Gemini 4 Carbon	Gemini 4 Argon	GPT-6 Astra	Claude Opus 5.5	Claude Fable 5.1
General capabilities	AA Index v5.0		59	50	57	58	49
	Vals Index v2.2		70.4%	69.2%	65.0%	70.2%	—
Agentic coding	SWE-sweep		9.6%	4.3%	9.5%	3.6%	2.9%
	FrontierSWE 2		61.2%	55.0%	65.5%	62.3%	56.3%
	Terminal-Bench 4.0 (AA)		65.2%	57.1%	59.1%	59.6%	52.0%
Computer use	Agents' Last Exam v1.1	Score	77.4%	69.0%	73.1%	71.6%	—
	OSWorld-2.1	Offline subset	81.5%	74.7%	80.9%	87.2%	76.8%
Knowledge work	AutomationBench	Zapier	53.9%	51.3%	41.4%	42.5%	31.4%
	Vending-Bench 2	Bank balance ($)	$16,484	$13,718	$15,515	$9,235	$5,422
Multimodal understanding	Chartography		72.5%	71.6%	71.0%	66.3%	46.2%
	Blueprint-Bench 2	Similarity score	58.6%	54.4%	49.7%	51.2%	41.9%
Long context	GraphWalks	256k to 1M, BFS (F1)	96.2%	84.2%	71.8%	66.8%	65.0%
Science and math	RiemannBench		80.0%	76.0%	72.0%	69.6%	65.6%
	Terminal-Bench-Science 0.1 (AA)		64.8%	52.9%	63.3%	59.0%	43.3%
Cybersecurity	CWE-bench		71.0%	68.0%	68.0%	67.0%	58.0%
	ExploitGym		54.8%	42.0%	42.4%	34.5%	—
	SEC-bench Pro		88.2%	83.4%	85.4%	—	—

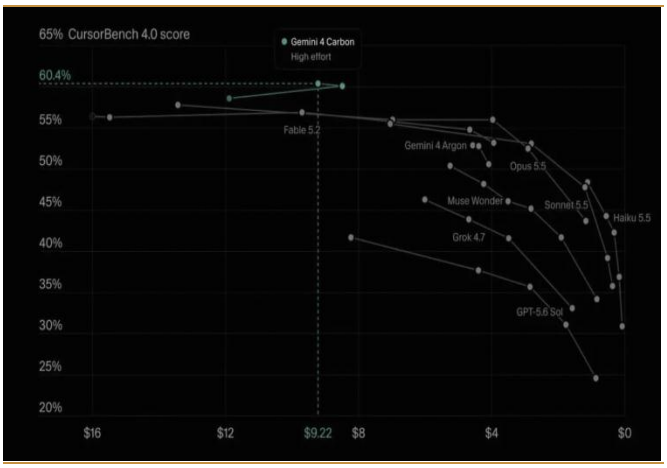
数据来源：Google Deepmind，西南证券整理

表 1：Google 部分模型定价

模型	输入价格（$/MTok）	输出价格（$/MTok）
Gemini 4 Carbon（首发促销期）	$2.00	$10.00
Gemini 4 Carbon（标准定价，12 月 1 日起）	$4.00	$20.00
Gemini 4 Argon（促销期 ）	$2.00	$10.00
Gemini 4 Argon（标准定价，12 月 1 日起）	$4.00	$20.00
Gemini 3.8 Flash（促销期）	$0.75	$3.75

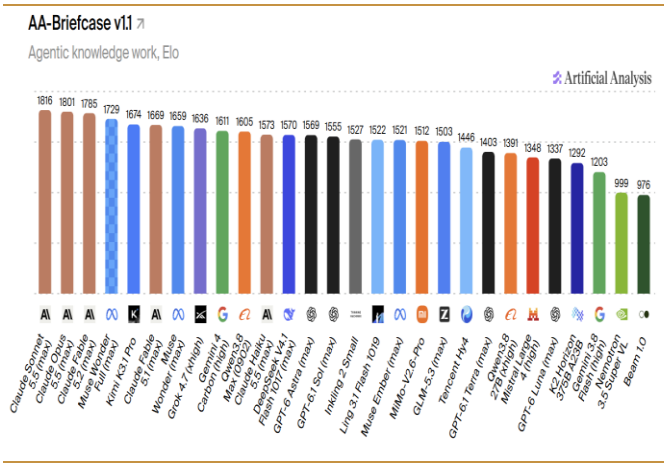
数据来源：Google，西南证券整理

图 2：CursorBench 4.0：各模型代码能力与单任务成本对比



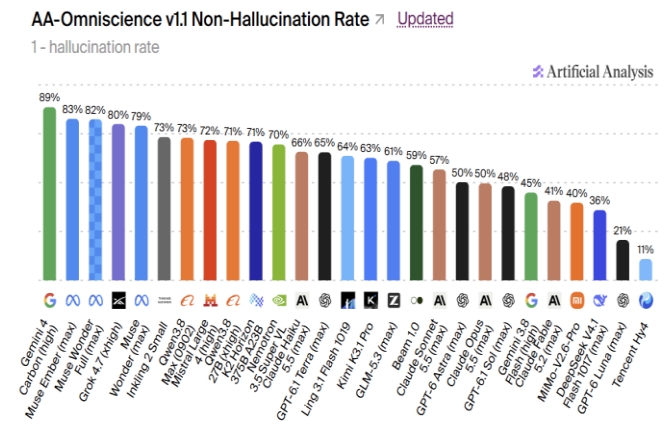
数据来源：Google DeepMind，CursorBench，西南证券整理

图 3：AA- Briefcase v1.1：各模型智能体知识工作 Elo 得分对比



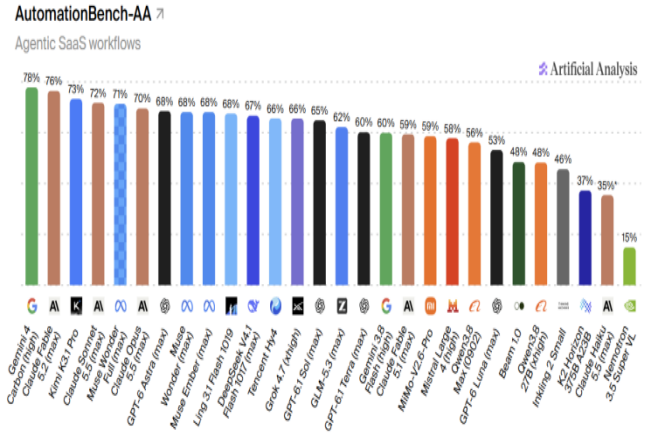
数据来源：Artificial Analysis，西南证券整理

图 4：AA- Omniscience v1.1 Non-Hallucination Rate：各模型非幻觉率对比



数据来源：Artificial Analysis，西南证券整理

图 5：AutomationBench- AA：各模型 SaaS 智能体工作流任务成功率对比



数据来源：Artificial Analysis，西南证券整理

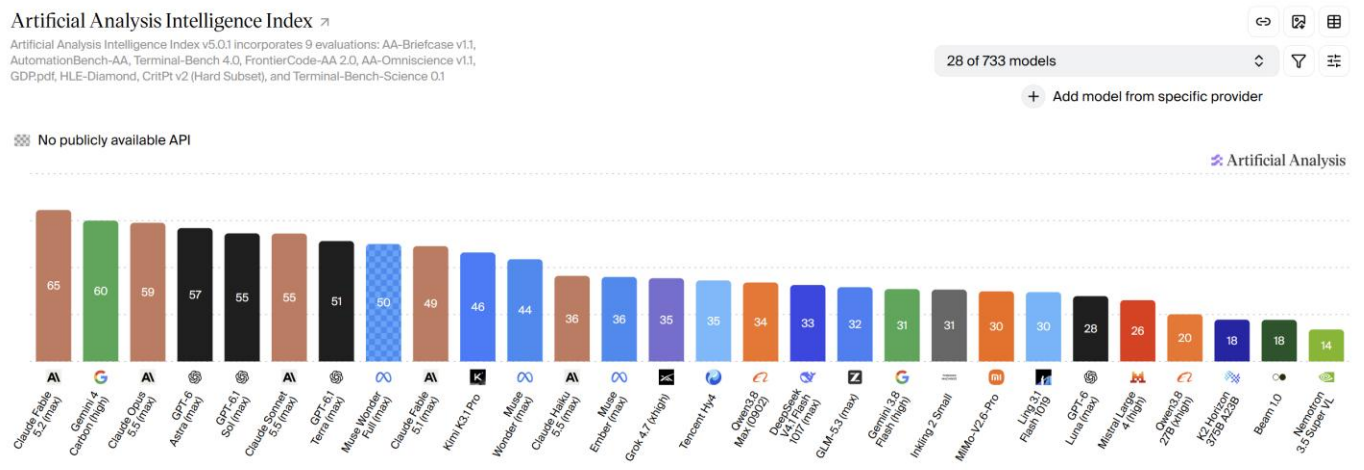
图 6：Claude Fable 5.2 与 Gemini 4 Carbon 多档位能力、成本、时延对比

Artificial Analysis

	Claude Fable 5.2 (max)	Claude Fable 5.2 (xhigh)	Claude Fable 5.2 (high)	Gemini 4 Carbon (high)	Gemini 4 Carbon (medium)	Claude Fable 5.2 (medium)	Gemini 4 Carbon (low)	Claude Fable 5.2 (low)
Intelligence								
Intelligence Index	65	62	62	60	58	56	54	54
Finance & Accounting	65	63	63	63	60	60	56	58
Strategy & Ops	67	65	64	61	58	62	53	61
Legal	63	62	60	62	59	58	54	57
Healthcare & Medical	65	--	--	57	--	--	--	--
Cost								
Cost per Task USD	$26.54	$14.83	$9.08	$7.76	$6.59	$3.80	$7.19	$3.10
Performance								
Time per Task Minutes	45.0	25.8	16.9	26.1	23.8	7.7	23.3	6.3
Time to First Answer Token s	247.85	69.21	30.02	66.01	23.21	12.88	12.24	8.14
Features								
Context Window	1M	1M	1M	1M	1M	1M	1M	1M
Creator	Anthropic	Anthropic	Anthropic	Google	Google	Anthropic	Google	Anthropic
License	Proprietary	Proprietary	Proprietary	Proprietary	Proprietary	Proprietary	Proprietary	Proprietary
Input Modality	T, Image, Audio, Video	T, Image, Audio, Video	T, Image, Audio, Video	T, Image, Audio, Video	T, Image, Audio, Video	T, Image, Audio, Video	T, Image, Audio, Video	T, Image, Audio, Video
Output Modality	T	T	T	T	T	T	T	T
Providers	↗	↗	↗	↗	↗	↗	↗	↗

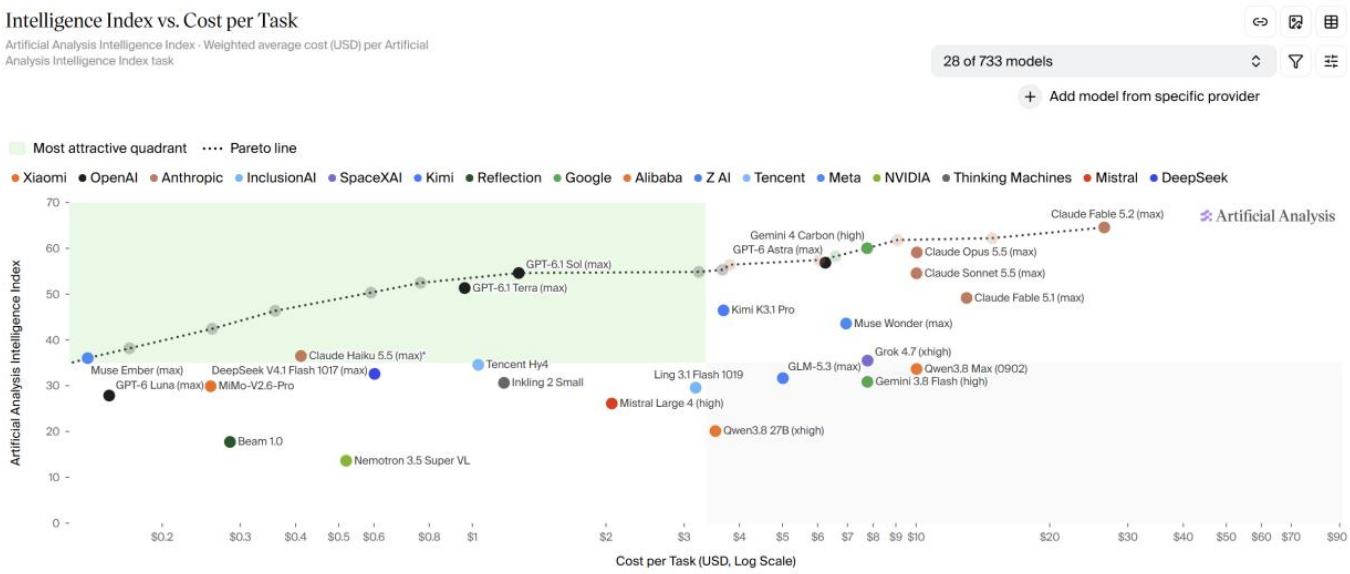
数据来源：Artificial Analysis，西南证券整理

图 7：Artificial Analysis 智能指数评分



数据来源：Artificial Analysis，西南证券整理

图 8：Artificial Analysis 大模型智能指数-成本散点图



数据来源：Artificial Analysis，西南证券整理

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，报告所采用的数据均来自合法合规渠道，分析逻辑基于分析师的职业理解，通过合理判断得出结论，独立、客观地出具本报告。分析师承诺不曾因，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获取任何形式的补偿。

投资评级说明

报告中投资建议所涉及的评级分为公司评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 个月内的相对市场表现，即：以报告发布日后 6 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。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；香港市场以恒生指数为基准；美国市场以纳斯达克综合指数或标普 500 指数为基准。

公司评级	买入：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅在 20%以上
	持有：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于 10%与 20%之间
	中性：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于-10%与 10%之间
	回避：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于-20%与-10%之间
	卖出：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅在-20%以下
行业评级	强于大市：未来 6 个月内，行业整体回报高于同期相关证券市场代表性指数 5%以上
	跟随大市：未来 6 个月内，行业整体回报介于同期相关证券市场代表性指数-5%与 5%之间
	弱于大市：未来 6 个月内，行业整体回报低于同期相关证券市场代表性指数-5%以下

重要声明

西南证券股份有限公司（以下简称“本公司”）具有中国证券监督管理委员会核准的证券投资咨询业务资格。

本公司与作者在自身所知情范围内，与本报告中所评价或推荐的证券不存在法律法规要求披露或采取限制、静默措施的利益冲突。

《证券期货投资者适当性管理办法》于 2017 年 7 月 1 日起正式实施，本报告仅供本公司签约客户使用，若您并非本公司签约客户，为控制投资风险，请取消接收、订阅或使用本报告中的任何信息。本公司也不会因接收人收到、阅读或关注自媒体推送本报告中的内容而视其为客户。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并进行交易，还可能为这些公司提供或争取提供投资银行或财务顾问服务。

本报告中的信息均来源于公开资料，本公司对这些信息的准确性、完整性或可靠性不作任何保证。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。

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。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，本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。

本报告及附录版权为西南证券所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。如引用须注明出处为“西南证券”，且不得对本报告及附录进行有悖原意的引用、删节和修改。未经授权刊载或者转发本报告及附录的，本公司将保留向其追究法律责任的权利。

西南证券研究院

上海

地址：上海市浦东新区陆家嘴 21 世纪大厦 10 楼

邮编：200120

北京

地址：北京市西城区金融大街 35 号国际企业大厦 A 座 8 楼

邮编：100033

深圳

地址：深圳市福田区益田路 6001 号太平金融大厦 22 楼

邮编：518038

重庆

地址：重庆市江北区金沙门路 32 号西南证券总部大楼 21 楼

邮编：400025

西南证券机构销售团队

区域	姓名	职务	手机	邮箱
上海	崔露文	上海公募销售主管	15642960315	clw@swsc.com.cn
	张方毅	私募销售主管	15821376156	zfyi@swsc.com.cn
	戴剑箫	销售岗	13524484975	daijx@swsc.com.cn
	叶佳缘	销售岗	15800609605	yejy@swsc.com.cn
	贾文婷	销售岗	13621609568	jiawent@swsc.com.cn
	毛玮琳	销售岗	18721786793	mwl@swsc.com.cn
	张大炜	销售岗	13163027178	zhangdaw@swsc.com.cn
	张怀文	销售岗	18800275611	zhanghuaiwen@swsc.com.cn
	苏湖	销售岗	18913127448	suhu1997@swsc.com.cn
北京	李杨	北京公募销售主管	18601139362	yfly@swsc.com.cn
	姚航	北京保险销售主管	15652026677	yhang@swsc.com.cn
	张岚	销售岗	18601241803	zhanglan@swsc.com.cn
	杨薇	销售岗	15652285702	yangwei@swsc.com.cn
	王宇飞	销售岗	18500981866	wangyuf@swsc.com.cn
	刘艳	销售岗	18456565475	liuyanyj@swsc.com.cn
	李逸凡	销售岗	13906354066	liyf@swsc.com.cn
广深	高欣	广深销售主管	13923418464	gaoxin@swsc.com.cn
	龚之涵	销售岗	15808001926	gongzh@swsc.com.cn
	文柳茜	销售岗	13750028702	wlq@swsc.com.cn
	林哲睿	销售岗	15602268757	lzr@swsc.com.cn
	黄诗洁	销售岗	18817316880	hsj@swsc.com.cn
	周玥	销售岗	18580531950	zhy1997@swsc.com.cn
	彭志伟	销售岗	15521604635	pzhw@swsc.com.cn