万联证券 WANLIAN SECURITIES

大模型调用效率提升，关注算力运营服务能力升级机遇

——计算机行业跟踪报告

同步大市(维持)

2026 年 11 月 16 日

行业核心观点：

上周，沪深300下跌0.94%，创业板指下跌1.59%，申万计算机行业指数下跌1.22%，分别跑输沪深300和跑赢创业板指0.28和0.37个百分点，在申万各一级行业中排名第15位。上周，Anthropic发布Claude Fable 5.2和Mythos 5.2，谷歌发布Gemini 4 Hydrogen。Fable 5.2在编码、推理、科学研究等任务上的表现有所提升，Anthropic称其API价格较前代下降20%、生成速度最高提升30%；Gemini 4 Hydrogen则面向日常编程和智能体工作流，并通过模型调度提升任务执行效率。我们认为，模型生成速度和执行效率提升，有助于缩短用户等待时间、提高单位时间内的任务处理能力，并改善多轮调用场景下的使用体验。随着智能体在编程、研究和办公等场景中持续运行，推理效率优化有望降低单位任务的时间与算力成本，提升高频调用的经济性，为智能体应用扩展及推理服务需求增长提供支撑。此外，智能体应用能否进一步普及，仍取决于其可靠性、安全性及包括人工监督在内的整体成本。IDC《全球半年度数据中心设施跟踪报告》预计全球数据中心机柜数量及设施资本支出将持续增长，电力、电网接入和制冷等条件将影响项目建设与算力供给。我们认为，AI应用发展将继续带动数据中心及配套设施需求，不同地区的电力条件、网络连接、客户结构和算力需求存在差异，单纯增加机柜和算力规模未必能同步提升实际使用效率，算力资源的利用效率和运营能力同样值得关注。建议关注智算中心的算力运营服务、算电协同以及液冷等基础设施的建设和运营环节。**中长期视角下，**继续聚焦AI产业、数据产业两大投资主线。

投资要点：

产业动态：（1）人工智能：IDC《全球半年度数据中心设施跟踪报告》数据显示，全球已配置机柜数量预计将从 2026 年的约 2500 万架增长至 2030 年的 5800 万架。**（2）AI 大模型：**Anthropic 于当地时间 11 月 9 日官宣推出 Claude Fable 5.2 和 Claude Mythos 5.2 两款前沿 AI 模型。**（3）AI 大模型：**谷歌于 11 月 12 日正式推出 Gemini 4 Hydrogen 模型，主打快速执行与高效运行，面向日常编码和智能体工作流场景。

行业估值：从估值情况来看，申万计算机行业 2026 年 11 月 13 日 PE-TTM 为 104.23 倍，低于 2023-2025 年历史 PE-TTM 的均值 158.58 倍。

风险提示：中美科技摩擦；地缘政治风险；产业生态建设不及预期；应用落地不及预期；数据隐私安全风险；市场竞争加剧。

行业相对沪深 300 指数表现



数据来源：聚源，万联证券研究所

相关研究

智能体终端管理规范落地，安全治理同步完善

我国人工智能推理数据量首超训练数据量，AI 产业需求重心正向应用推理延伸

从 Anthropic 招股书看 AI 大模型商业化进程及算力需求

分析师： 夏清莹
执业证书编号：S0270520050001
电话：(0755) 8322 3620
邮箱：xiaqy1@wlzq.com.cn
研究助理： 梁源源
电话：13242092258
邮箱：liangyy@wlzq.com.cn

正文目录

1 行业周观点.....3
2 市场行情回顾.....3
3 产业动态.....4
3.1 人工智能：IDC 预计 2030 年，全球数据中心机柜量将达 5800 万架.....4
3.2 AI 大模型：Anthropic 最强 Claude Fable/Mythos 5.2 AI 模型登场，API 调用费下降 20%.....5
3.3 AI 大模型：谷歌发布 Gemini 4 Hydrogen 模型：主打日常编程与智能体工作流，百万 Token 输入 1.5 美元.....6
4 风险提示.....6

图表 1： 沪深 300、创业板指、申万一级行业指数涨跌幅（%）（上周）.....3
图表 2： 沪深 300、创业板指、申万一级行业指数涨跌幅（%）（2026 年初至 2026 年 11 月 13 日）.....4
图表 3： 申万计算机行业估值情况（PE-TTM，单位：倍）.....4

1 行业周观点

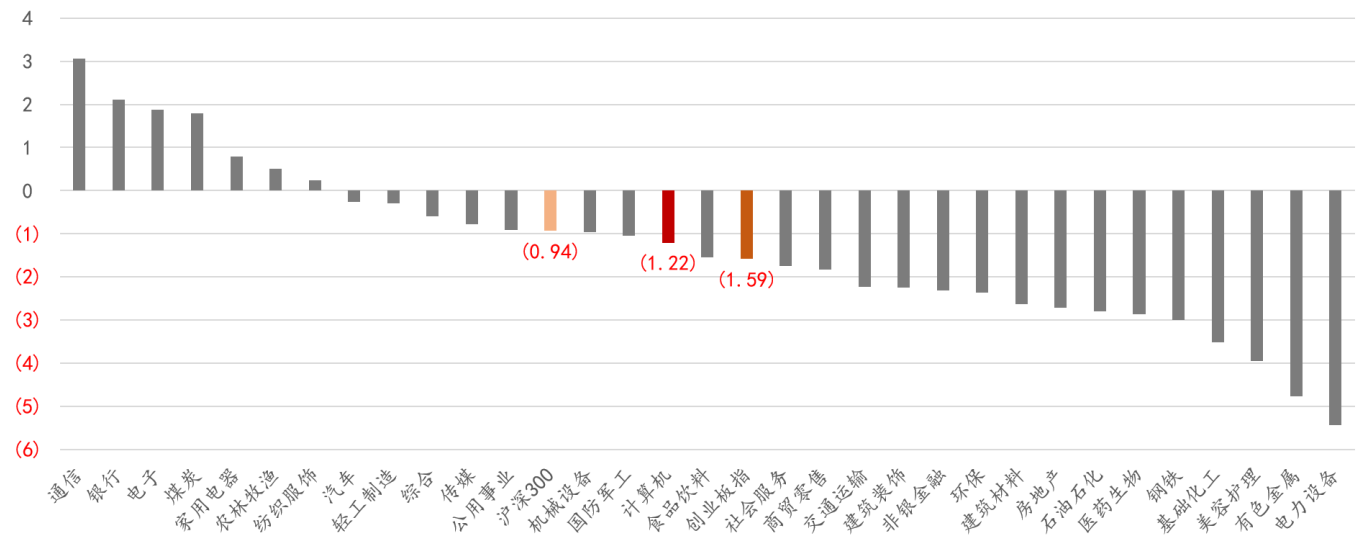
上周，Anthropic发布Claude Fable 5.2和Mythos 5.2，谷歌发布Gemini 4 Hydrogen。Fable 5.2在编码、推理、科学研究等任务上的表现有所提升，Anthropic称其API价格较前代下降20%、生成速度最高提升30%；Gemini 4 Hydrogen则面向日常编程和智能体工作流，并通过模型调度提升任务执行效率。我们认为，模型生成速度和执行效率提升，有助于缩短用户等待时间、提高单位时间内的任务处理能力，并改善多轮调用场景下的使用体验。随着智能体在编程、研究和办公等场景中持续运行，推理效率优化有望降低单位任务的时间与算力成本，提升高频调用的经济性，为智能体应用扩展及推理服务需求增长提供支撑。此外，智能体应用能否进一步普及，仍取决于其可靠性、安全性及包括人工监督在内的整体成本。IDC《全球半年度数据中心设施跟踪报告》预计全球数据中心机柜数量及设施资本支出将持续增长，电力、电网接入和制冷等条件将影响项目建设与算力供给。我们认为，AI应用发展将继续带动数据中心及配套设施需求，不同地区的电力条件、网络连接、客户结构和算力需求存在差异，单纯增加机柜和算力规模未必能同步提升实际使用效率，算力资源的利用效率和运营能力同样值得关注。建议关注智算中心的算力运营服务、算电协同以及液冷等基础设施的建设和运营环节。

中长期视角下，继续聚焦AI产业、数据产业两大投资主线。1）AI产业：主要关注国内AI大模型的技术升级和产品迭代、头部AI大模型厂商的生态建设、智能驾驶、具身智能等新型智能终端及智能体的应用落地进程，以及人工智能治理的生态建设。2）数据产业：主要关注数据基础设施建设、高质量数据集等数据要素资源的开发利用、数据要素的价值释放及应用落地以及数据隐私安全。

2 市场行情回顾

申万计算机行业分别跑输沪深300和跑赢创业板指，周涨跌幅位于申万31个一级行业的第15名。上周，沪深300下跌0.94%，创业板指下跌1.59%，申万计算机行业指数下跌1.22%，分别跑输沪深300和跑赢创业板指0.28和0.37个百分点，在申万各一级行业中排名第15位。

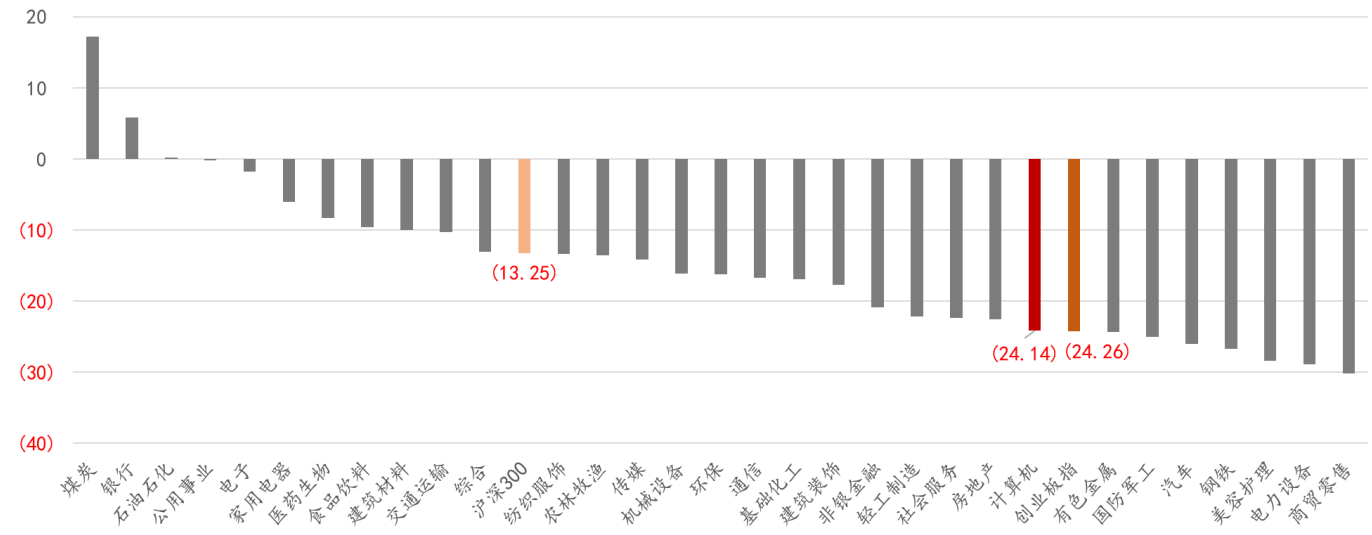
图表1：沪深300、创业板指、申万一级行业指数涨跌幅（%）（上周）



资料来源：iFinD，万联证券研究所

申万计算机行业2026年初至11月13日累计下跌24.14%，位于申万31个一级行业的第24名。2026年初至11月13日，沪深300下跌13.25%，创业板指下跌24.26%，申万计算机行业指数下跌24.14%，分别跑输沪深300和跑赢创业板指10.89和0.12个百分点，在申万各一级行业中排名第24位。

图表2：沪深300、创业板指、申万一级行业指数涨跌幅（%）（2026年初至2026年11月13日）



资料来源：iFinD，万联证券研究所

行业估值水平低于历史中枢水平。从估值情况来看，申万计算机行业2026年11月13日PE-TTM为104.23倍，低于2023-2025年历史PE-TTM的均值158.58倍。

图表3：申万计算机行业估值情况（PE-TTM，单位：倍）



资料来源：iFinD，万联证券研究所

3 产业动态

3.1 人工智能：IDC预计2030年，全球数据中心机柜量将达5800万架

11月11日消息，IDC《全球半年度数据中心设施跟踪报告》数据显示，全球已配置机柜数量预计将从2026年的约2500万架增长至2030年的5800万架。

同期，全球数据中心年度能源消耗预计将从2025年的519太瓦时（TWh）增长至2030年

的1308太瓦时；全球数据中心设施资本支出预计将从2025年的2,760亿美元（现汇率约合1.85万亿元人民币）增长至2030年的6,550亿美元（现汇率约合4.38万亿元人民币）。

该统计不含服务器、GPU等IT基础设施投入，主要反映电力、制冷、建筑等物理设施的建设需求。

AI基础设施扩张不仅是计算设备规模的增长，更需要庞大的物理设施与能源资源支撑。但设施供给增长并非总能与需求同步：电力供应、电网接入与数据中心建设能力，正成为全球部分市场近期增长的重要制约因素，关键电力设备、液冷及相关供应链的交付周期也在影响新增设施的建设进度。

全球数据中心供给面临约束，并不代表所有地区都处于供不应求状态。IDC全球研究指出，部分地区资源紧缺与其他地区容量利用不足可能同时存在，AI基础设施面临的不仅是总体供给规模问题，更包括不同市场之间的需求与资源匹配问题。

IDC指出，这一观察对中国市场尤为重要：根据工业和信息化部2026年7月披露的数据，截至6月底，中国智能算力规模已达到2185 EFLOPS（FP16），同比增长177%，全国算力设施整体上架率为71.4%，反映出中国算力基础设施规模持续增长的同时，资源利用仍是观察市场发展的重要维度。

中国数据中心市场的供给主体具有自身特点，三大电信运营商、第三方数据中心设施服务商与超大规模云服务商共同构成核心供给力量，不同类型企业拥有差异化的客户基础、业务模式与设施布局。

与此同时，AI基础设施的需求主体也在发生变化：除传统超大规模云服务商外，专业AI云服务商（Neocloud）、前沿模型厂商以及具有大规模AI需求的企业，正成为新的基础设施采购力量，AI工作负载也正从集中式云环境，向私有环境、边缘及专用AI基础设施等多种部署形态扩展。

这意味着中国市场未来的设施需求增长，将通过不同类型的服务提供商与客户群体逐步释放，行业需要同时关注需求来源、服务提供方式与运营商竞争结构的变化。（资料来源：IT之家）

3.2 AI 大模型：Anthropic 最强 Claude Fable/Mythos 5.2 AI 模型登场，API 调用费下降 20%

11月10日消息，Anthropic于当地时间11月9日官宣推出Claude Fable 5.2和Claude Mythos 5.2，两款模型基于同一底层模型，前者面向普通用户并内置网络安全与生物安全防护，后者则面向通过审核的网络安全和生命科学机构开放，用于高风险的专业应用。官方称，新模型在编码、推理、科学研究、网络安全及机器人等领域均有显著提升，同时输出速度最高提高30%，成本下降20%。

定价方面，Claude Fable 5.2每100万词元（Token）输入费用为8美元（现汇率约合53.4元人民币），每100万词元（Token）输出费用为40美元（现汇率约合267.2元人民币）。作为对比，Claude Fable 5.1的每100万词元输入/输出价格分别为10美元（现汇率约合66.8元人民币）和50美元（现汇率约合334元人民币），按此计算5.2版本的API调用费用下调20%。

生成速度方面，官方表示相比Fable 5.1，最新Claude Fable 5.2模型生成速度最高提升30%，而且在可选“快速模式”下，生成速度最高再提升至2.5倍，但价格相应更高。

在预发布测试中，Fable 5.2完成了多项引人注目的任务：设计出三款可实际运行的电子设备，包括语音助手中枢、USB会议音箱和手持虚拟宠物；协调逾1,000个子智能体，在8天内从零构建出可在笔记本电脑启动的原型操作系统；并以约1,000个代理在30小时内完成费马大定理的Lean形式化证明，远快于此前11天的纪录。（资料来源：IT之家）

3.3 AI 大模型：谷歌发布 Gemini 4 Hydrogen 模型：主打日常编程与智能体工作流，百万 Token 输入 1.5 美元

11月13日消息，谷歌于11月12日正式推出Gemini 4 Hydrogen模型，主打快速执行与高效运行，面向日常编码和智能体工作流场景，将Gemini 4系列能力全面下放至所有层级用户。

输入：每百万Tokens 1.50美元（现汇率约合10.0元人民币）

输出：每百万Tokens 7.50美元（现汇率约合50.1元人民币）

该定价覆盖全100万Token上下文窗口，无分段阶梯计价。在Artificial Analysis智能指数中，Hydrogen默认高思考档位得分51，高于Gemini 4 Argon的50，最高xhigh档位可达54，同时任务完成速度常优于同级别模型。

谷歌同步推出Orchestrator智能调度功能，内置于Google Antigravity平台，可在同一会话中自动在Carbon与Hydrogen等模型间动态切换。在Terminal-Bench 4.0测试中，该调度方案达到接近Carbon的任务成功率，同时平均消耗配额减少24%，完成速度提升19%。

对比上一代Gemini 3.8 Flash，Hydrogen默认高思考档位下任务完成速度提升2.5 倍，Terminal-Bench 4.1得分从18.2%提升至57%，虚构错误率降低5倍以上。

该模型设置四档思考深度：low、medium、high（默认）、xhigh，可根据任务需求灵活调节推理深度、延迟与成本；支持最高100万Token原生输入、100万Token输出，可处理大型PDF、表格、图表、音视频等多模态内容。

安全方面，Hydrogen将智能体任务越权行为减少最高81%，服务不足问题减少79%，工具调用与编码场景下的失准行为、任务走捷径问题减少12倍；内部智能体诚实度评测得分96.7%，思维链保留率达99.9%，在Gray Swan IPI基准上对抗间接提示注入的韧性提升2.3倍以上。（资料来源：IT之家）

4 风险提示

中美科技摩擦；地缘政治风险；产业生态建设不及预期；应用落地不及预期；数据隐私安全风险；市场竞争加剧。

行业投资评级

强于大市：未来6个月内行业指数相对大盘涨幅10％以上；
同步大市：未来6个月内行业指数相对大盘涨幅10％至-10％之间；
弱于大市：未来6个月内行业指数相对大盘跌幅10％以上。

公司投资评级

买入：未来6个月内公司相对大盘涨幅15％以上；
增持：未来6个月内公司相对大盘涨幅5％至15％；
观望：未来6个月内公司相对大盘涨幅-5％至5％；
卖出：未来6个月内公司相对大盘跌幅5％以上。
基准指数 ： 沪深300指数

风险提示

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。

证券分析师承诺

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，以勤勉的执业态度，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。

免责条款

万联证券股份有限公司（以下简称“本公司”）是一家覆盖证券经纪、投资银行、投资管理和证券咨询等多项业务的全国性综合类证券公司。本公司具有中国证监会许可的证券投资咨询业务资格。
本报告仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。本报告中的信息或所表述的意见并未考虑到个别投资者的具体投资目的、财务状况以及特定需求。客户应自主作出投资决策并自行承担投资风险。本公司不对任何人因使用本报告中的内容所导致的损失负任何责任。在法律许可情况下，本公司或其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、财务顾问或类似的金融服务。
市场有风险，投资需谨慎。本报告是基于本公司认为可靠且已公开的信息撰写，本公司力求但不保证这些信息的准确性及完整性，也不保证文中的观点或陈述不会发生任何变更。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。分析师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。
本报告的版权仅为本公司所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登、发表和引用。未经我方许可而引用、刊发或转载的引起法律后果和造成我公司经济损失的概由对方承担，我公司保留追究的权利。

万联证券股份有限公司 研究所
上海浦东新区世纪大道 1528 号陆家嘴基金大厦
北京西城区平安里西大街 28 号中海国际中心
深圳福田区深南大道 2007 号金地中心
广州天河区珠江东路 11 号高德置地广场