

计算机

计算机行业周报（20250818-20250824）： Deepseek 引领国产 AI 产业发展

■ 走势比较



■ 推荐公司及评级

相关研究报告

<<业绩恢复快速增长, AI 融合产品应用>>—2025-08-25

<<订单增速大幅回暖, 海外业务稳步推进>>—2025-08-25

<<业绩持续高增, AI 全面赋能>>—2025-08-25

证券分析师: 曹佩

电话:

E-MAIL: caopeisz@tpyzq.com

分析师登记编号: S1190520080001

证券分析师: 王景宜

电话:

E-MAIL: wangjy@tpyzq.com

分析师登记编号: S1190523090002

报告摘要

行情回顾: 本周（2025. 8. 18-2025. 8. 24）申万计算机行业指数上涨 7. 93%，在 31 个申万一级行业中排名第 4 位。截至本周五（8 月 22 日），计算机行业指数与较年初上涨 28. 86%，跑赢沪深 300 指数 17. 60pct。

DeepSeek-V3. 1 正式发布, 推理效率提升, Agent 能力增强。8 月 21 日, 深度求索发布并开源 DeepSeek-V3. 1。本次升级包含以下主要变化: 1) 混合推理架构: 一个模型同时支持思考模式与非思考模式, 且上下文均已扩展为 128K; 2) 更高的思考效率: 相比 DeepSeek-R1-0528, DeepSeek-V3. 1-Think 在保持同等回答质量的情况下响应速度更快, 同时成绩基本持平; 3) 更强的 Agent 能力: 通过 Post-Training 优化, 新模型在工具使用与智能体任务中的表现有较大提升, 编程和搜索智能体性能显著增强。

编程和搜索智能体能力显著增强。DeepSeek 表示, DeepSeek-V3. 1 通过后训练优化, 在工具使用和多步骤代理任务上实现了显著提升, 是迈向 Agent (智能体) 时代的第一步。在编程智能体方面, DeepSeek-V3. 1 在代码修复测评 SWE 与命令行终端环境下的复杂任务 (Terminal-Bench) 测试中, 相比之前的 DeepSeek 系列模型有明显提高。在搜索智能体方面, DeepSeek-V3. 1 为思考模式中的搜索工具调用设计了专门格式, 支持复杂的多轮搜索代理任务。V3. 1 在需要多步推理的复杂搜索测试 (browsecomp) 上大幅领先 R1-0528, 在多学科专家级难题测试 (HLE, 基于仅文本子集) 上性能显著提升, 并且能够有效支持访问外部或最新信息的复杂问题处理。

模型 API 定价性价比优势依然突出。价格方面, V3. 1 输入价格为 0. 5 元/百万 tokens (缓存命中), 4 元/百万 tokens (缓存未命中); 输出价格为 12 元/百万 tokens。根据腾讯科技, V3. 1 完成一次完整的编程任务的成本仅需约 1. 01 美元, 远低于 Claude 4 Opus (便宜 68 倍), 在各主流模型中性价比名列前茅。

DeepSeek 为下一代国产芯片设计 UE8M0 FP8, 支持国产算力生态。DeepSeek-V3. 1 创新采用了专为 AI 计算优化的数值格式 UE8M0 FP8, 通过简化计算、节省显存和带宽, 大幅提升推理效率和降低部署成本。DeepSeek 官方明确指出 UE8M0 FP8 精度格式是针对即将发布的下一代国产芯片设计。根据量子位, 目前寒武纪、海光信息、摩尔线程、沐曦等厂商均推出了支持 FP8 的算力芯片, 而昇腾 910B 和 910C 暂不支持原生 FP8, 但官方路线图已经写明“2025Q4 原生 FP8”。

投资建议：Deepseek 新版本的发布首先会体现对国产算力产业的拉动，重点推荐海光信息、智微智能、联想集团等。应用端重点推荐金山办公、用友网络、赛意信息、熵基科技、云从科技等。

风险提示：AI 技术发展不及预期，市场竞争加剧。

目录

一、	DeepSeek-V3.1 正式发布，引领国产 AI 产业发展	5
二、	行情回顾	8
三、	投资建议	8
四、	风险提示	9

图表目录

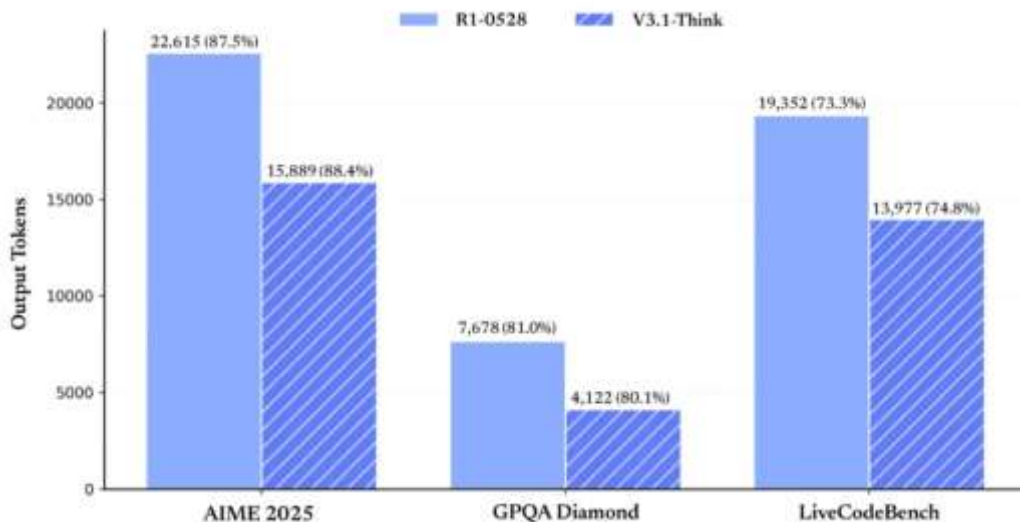
图表 1: R1-0528 与 V3.1-Think 的 token 消耗量对比图	5
图表 2: DeepSeek-V3.1 编程智能体测评	6
图表 3: DeepSeek-V3.1 API 定价	7
图表 4: 支持 FP8 计算精度的国产算力芯片	7
图表 5: 申万一级行业周涨幅 (8.18-8.22)	8
图表 6: 申万计算机指数和沪深 300 指数年初至今涨跌幅 (截至 8.22)	8
图表 7: 本周计算机行业涨跌幅前五	8

一、DeepSeek-V3.1 正式发布，引领国产 AI 产业发展

8月21日，深度求索发布并开源 DeepSeek-V3.1。本次升级包含以下主要变化：1) 混合推理架构：一个模型同时支持思考模式与非思考模式，且上下文均已扩展为 128K；2) 更高的思考效率：相比 DeepSeek-R1-0528，DeepSeek-V3.1-Think 能在更短时间内给出答案；3) 更强的 Agent 能力：通过 Post-Training 优化，新模型在工具使用与智能体任务中的表现有较大提升。

模型思考效率明显提升。官方测试数据显示，新版本 V3.1-Think 比起相较于 R1-0528，在保持同等回答质量的情况下响应速度更快，同时成绩基本持平。V 3.1 在 AIME 2025 的得分为 88.4 分 (R1-0528 为 87.5 分)，GPQA 为 80.01 分 (R1-0528 为 81 分)，liveCodeBench 为 74.8 分 (R1-0528 为 73.3 分)。成绩基本持平的情况下，V3.1-Think 输出 token 数减少了 20%-50%。同时，V3.1 在非思考模式下的输出长度也得到了有效控制，相比于 DeepSeek-V3-0324，能够在输出长度明显减少的情况下保持相同的模型性能。

图表1: R1-0528 与 V3.1-Think 的 token 消耗量对比图



资料来源: DeepSeek 微信公众号, 太平洋证券整理

编程和搜索智能体能力显著增强。DeepSeek 表示，DeepSeek-V3.1 通过后训练优化，在工具使用和多步骤代理任务上实现了显著提升，是迈向 Agent（智能体）时代的第一步。在编程智能体方面，DeepSeek-V3.1 在代码修复测评 SWE 与命令行终端环境下的复杂任务 (Terminal-Bench) 测试中，相比之前的 DeepSeek 系列模型有明显提高。在搜索智能体方面，DeepSeek-V3.1 为思考

模式中的搜索工具调用设计了专门格式，支持复杂的多轮搜索代理任务。V3.1 在需要多步推理的复杂搜索测试 (browsecomp) 上大幅领先 R1-0528，在多学科专家级难题测试 (HLE，基于仅文本子集) 上性能显著提升，并且能够有效支持访问外部或最新信息的复杂问题处理。

图表2: DeepSeek-V3.1 编程智能体测评

Benchmarks	DeepSeek-V3.1	DeepSeek-V3-0324	DeepSeek-R1-0528
SWE-bench Verified	66.0	45.4	44.6
SWE-bench Multilingual	54.5	29.3	30.5
Terminal-Bench	31.3	13.3	5.7

资料来源: DeepSeek 微信公众号, 太平洋证券整理

API 定价性价比优势依然突出。价格方面, V3.1 输入价格为 0.5 元/百万 tokens (缓存命中), 4 元/百万 tokens (缓存未命中); 输出价格为 12 元/百万 tokens。根据腾讯科技, V3.1 完成一次完整的编程任务的成本仅需约 1.01 美元, 远低于 Claude 4 Opus (便宜 68 倍), 在各主流模型中性价比名列前茅。

图表3: DeepSeek-V3.1 API 定价



资料来源: DeepSeek 微信公众号, 太平洋证券整理

DeepSeek-V3.1 创新采用 UE8M0 FP8 精度格式, 具备更小带宽、更低功耗和更高吞吐的优势。

DeepSeek-V3.1 的采用了 UE8M0 FP8 Scale 参数精度。根据凤凰网科技, FP8 是一种将数字压缩到仅用 8 位 (bit) 来存储的浮点格式, 旨在用更少的数据位宽降低 AI 大模型训练和推理中对显存带宽的空前压力。FP8 相对于传统的 FP16 计算能够实现两倍的浮点算力提升、访存和通信带宽效率提升和存储容量利用率提升, 同时最优化张量表达精度。

DeepSeek 明确支持国产算力生态。DeepSeek 官方明确指出 UE8M0 FP8 精度格式是针对即将发布的下一代国产芯片设计。根据量子位, 目前寒武纪、海光信息、摩尔线程、沐曦等厂商均推出了支持 FP8 的算力芯片, 而昇腾 910B 和 910C 暂不支持原生 FP8, 但官方路线图已经写明“2025Q4 原生 FP8”。

图表4: 支持 FP8 计算精度的国产算力芯片

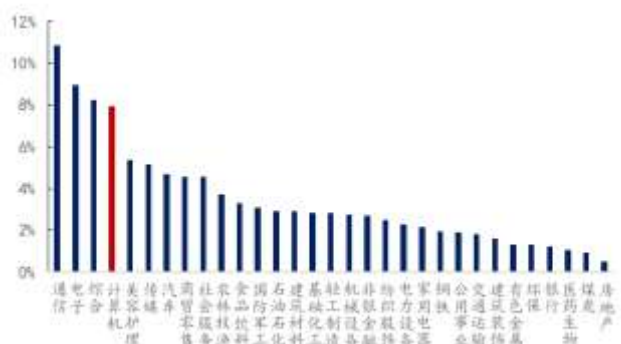
厂商	对 FP8 精度支持情况
寒武纪	公司旗下的 MLU370-S4、思元 590 及最新 690 系列芯片均支持 FP8 计算
海光信息	深算三号 DCU 支持 FP8 计算, 存在进一步优化的空间
华为昇腾	官方路线图写明 2025Q4 支持原生 FP8
中昊芯英	“刹那” TPU AI 芯片支持 FP8 精度
摩尔线程	旗舰产品 MTT S5000 支持 FP8 精度计算
沐曦	公司旗下的曦云 C600 支持 FP8 精度计算

资料来源: 量子位, 太平洋证券整理

二、行情回顾

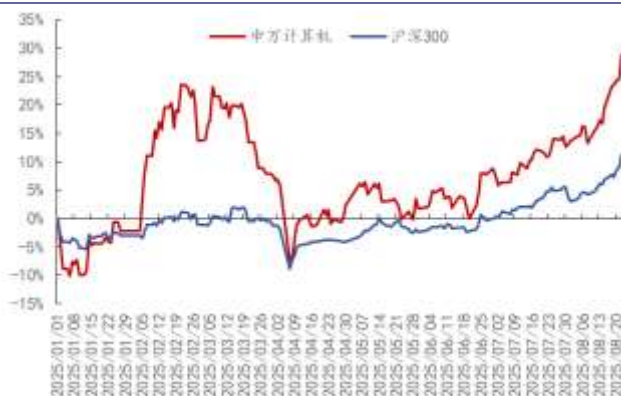
本周申万计算机行业指数上涨 7.93%, 在 31 个申万一级行业中排名第 4 位。截至本周五 (8 月 22 日), 计算机行业指数与较年初上涨 28.86%, 跑赢沪深 300 指数 17.60pct。

图表5: 申万一级行业周涨幅 (8.18-8.22)



资料来源: iFinD, 太平洋证券整理

图表6: 申万计算机指数和沪深 300 指数年初至今涨跌幅 (截至 8.22)



资料来源: iFinD, 太平洋证券整理

本周计算机行业涨幅前五的公司分别御银股份 (+46.2%)、诚迈科技 (+34.92%)、科创信息 (+33.94%)、中科金财 (+33.55%)、指南针 (+28.63%); 跌幅前五的公司分别为佳缘科技 (-10.73%)、思创医惠 (-10.08%)、运达科技 (-7.76%)、汇金股份 (-7.43%)、莱斯信息 (-6.79%)。

图表7: 本周计算机行业涨跌幅前五

涨幅前五			跌幅前五		
证券简称	周涨跌幅 (%)	总市值 (亿元)	证券简称	周涨跌幅 (%)	总市值 (亿元)
御银股份	46.20	76.12	佳缘科技	-10.73	47.16
诚迈科技	34.92	128.78	思创医惠	-10.08	39.91
科创信息	33.94	57.20	运达科技	-7.76	59.13
中科金财	33.55	160.40	汇金股份	-7.43	81.77
指南针	28.63	859.50	莱斯信息	-6.79	155.97

资料来源: iFinD, 太平洋证券整理

三、投资建议

Deepseek 新版本的发布首先会体现对国产算力产业的拉动, 重点推荐海光信息、智微智能、联想集团等。应用端重点推荐金山办公、用友网络、赛意信息、熵基科技、云从科技等。

四、风险提示

AI 技术发展不及预期，市场竞争加剧。

投资评级说明

1、行业评级

看好: 预计未来 6 个月内, 行业整体回报高于沪深 300 指数 5%以上;

中性: 预计未来 6 个月内, 行业整体回报介于沪深 300 指数-5%与 5%之间;

看淡: 预计未来 6 个月内, 行业整体回报低于沪深 300 指数 5%以下。

2、公司评级

买入: 预计未来 6 个月内, 个股相对沪深 300 指数涨幅在 15%以上;

增持: 预计未来 6 个月内, 个股相对沪深 300 指数涨幅介于 5%与 15%之间;

持有: 预计未来 6 个月内, 个股相对沪深 300 指数涨幅介于-5%与 5%之间;

减持: 预计未来 6 个月内, 个股相对沪深 300 指数涨幅介于-5%与-15%之间;

卖出: 预计未来 6 个月内, 个股相对沪深 300 指数涨幅低于-15%以下。

太平洋证券股份有限公司

云南省昆明市盘龙区北京路 926 号同德广场写字楼 31 楼



研究院

中国北京 100044

北京市西城区北展北街九号

华远·企业号 D 座

投诉电话： 95397

投诉邮箱： kefu@tpyzq.com

免责声明

太平洋证券股份有限公司（以下简称“我公司”或“太平洋证券”）具备中国证券监督管理委员会核准的证券投资咨询业务资格。

本报告仅向与太平洋证券签署服务协议的签约客户发布，为太平洋证券签约客户的专属研究产品，若您并非太平洋证券签约客户，请取消接收、订阅或使用本报告中的任何信息；太平洋证券不会因接收人收到、阅读或关注媒体推送本报告中的内容而视其为太平洋证券的客户。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机构和个人的投资建议，投资者应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。

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映研究人员的个人观点。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员对使用本报告及其内容所引发的任何直接或间接损失概不负责。我公司或关联机构可能会持有报告中所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业务服务。本报告版权归太平洋证券股份有限公司所有，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。任何人使用本报告，视为同意以上声明。