

# 计算机行业深度报告

## 国产算力趋势不可逆，AI 信创的战略拐点

增持（维持）

2026 年 05 月 09 日

证券分析师 王紫敬

执业证书：S0600521080005

021-60199781

wangzj@dwzq.com.cn

### 投资要点

■ **智能体驱动的 Token 经济革命，国产替代开启战略机遇期。2026 年以来，AI 产业商业价值迎来结构性裂变。**智能体全面商业化落地，推动 Token 经济爆发式增长，并重构 AI 使用范式。相比传统 Chatbot 单轮问答仅消耗数百 Token，企业级智能体完成一次复杂任务往往需要数十万至数百万 Token，直接带动国内 Token 调用量高速增长。截至 2026 年 3 月，我国日均 Token 调用量已超过 140 万亿，较 2024 年初增长 1000 多倍；2026 年 4 月最后一周，中国 AI 大模型周调用量升至 7.942 万亿 Token，环比增长 81.7%，并再次超过美国。Token 正从技术计量单位演变为 AI 时代核心生产资料和数字硬通货，全球 AI 竞争焦点也从模型参数堆砌转向 Token 生态构建，算力需求重心由训练主导转向推理主导。

■ **DeepSeek V4 首次使用国产算力参与训练，标志着 AI 信创进入战略机遇期，也成为国产算力由政策驱动走向产业自证的重要拐点。**这一事件突破了市场对国产算力仅适用于政务信创、推理部署和中低端训练场景的固有认知，说明国产算力正在进入高强度、大规模复杂模型训练环节，并推动中国 AI 产业链向“国产模型+国产芯片”的训推闭环演进。在海外对华 AI 芯片、模型能力和技术交流限制持续高频化、体系化、长期化的背景下，AI 信创已不再是传统 IT 国产化的延伸，而是国家级科技竞争中的底层基础设施话语权。

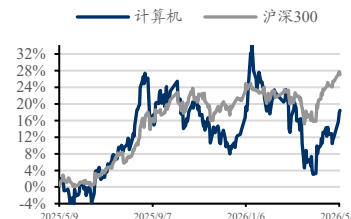
■ **AI 信创产业正形成五大核心主线：GPU 芯片、CPU 芯片、昇腾产业链、算力租赁和信创大模型，覆盖从底层硬件到上层应用的完整链条。**国内算力基础设施建设已基本完成甚至超额完成“十四五”规划目标，“十五五”期间，核心技术自主可控与算力基础设施建设仍将是政策、资本开支和产业落地的共同主线。国产算力替代呈现出推理侧先行、训练侧突破、生态侧协同的特征。**GPU 芯片是 AI 信创最核心、最直接的受益方向，是大模型训练和推理的关键底座。**DeepSeek V4 采用国产算力参与训练，意味着国产 GPU 及 AI 加速芯片已从推理侧验证迈向训练侧验证，进入真实大模型训练场景后，产业迭代有望明显提速。**CPU 芯片则在 Agent 时代成为底层调度核心。**随着 AI 算力需求高速增长、通用服务器更新周期启动，以及 Intel、AMD 相继提价，服务器 CPU 供需错配加剧。Agent 与强化学习重构 AI 工作负载范式，算力基础设施建设也从单纯“GPU 堆叠”转向“GPU+CPU 异构协同”，CPU 在 AI 基础架构中的地位持续提升。

■ **昇腾产业链、算力租赁和信创大模型则构成 AI 信创的重要增长极。**昇腾产业链具备体系化生态特征，覆盖材料、连接器、PCB、服务器整机、集成、算力平台和行业应用，未来价值有望从政策采购和行业信创项目扩展至通用 AI 基础设施市场。算力租赁是短期高弹性方向，在高端算力供需缺口显著、Agent 规模化落地的背景下，商业模式正从“裸算力出租”向参与大模型 Token 收入分成演进，估值逻辑也将由“租金收益型”转向“Token 分成型”。信创大模型则是尚未充分定价的核心赛道，党政机关、地方政府和央企对自主可控大模型需求持续释放，政务及重点行业对安全性、私有化部署和国产算力适配提出更高要求，“人工智能+教育”等国家级政策也提供了高确定性落地场景。

■ **投资建议：GPU：**寒武纪、海光信息、沐曦股份、摩尔线程；**CPU：**禾盛新材、海光信息、中国长城、龙芯中科；**昇腾：**云南锗业、华丰科技、航天电器、华正新材、神州数码；**算力租赁：**宏景科技、协创数据、盈峰环境、利通电子；**信创大模型：**科大讯飞、智谱、minimax。

■ **风险提示：**若政策落地不及预期、技术迭代不及预期、算力供给波动风险、商业化兑现风险、市场竞争加剧风险。

### 行业走势



### 相关研究

《国内算力时代开启：从算力租赁到国产算力》

2026-05-04

《国产化训练从 0 到 1 里程碑，战略意义大于性能意义》

2026-04-30

## 内容目录

<b>1. AI 商业价值爆发式增长：智能体驱动的 Token 经济革命，国产替代开启战略机遇期</b> .....	<b>4</b>
1.1. Token 消耗爆发.....	4
1.2. DeepSeek V4 国产算力参与训练：AI 信创的战略拐点.....	5
<b>2. AI 信创五大主线，国产替代战略价值显著抬升</b> .....	<b>6</b>
2.1. 国产算力：基础设施扩容与国产替代同步推进.....	6
2.1.1. GPU 芯片：国产算力替代的第一核心方向.....	7
2.1.2. CPU 芯片：Agent 时代的底层调度核心.....	7
2.2. 昇腾链：国产 AI 算力生态的体系化主线.....	8
2.3. 算力租赁：短期高弹性方向，商业模式迎来重估.....	8
2.4. 信创大模型：党政、地方政府与央国企需求尚未充分定价.....	9
<b>3. 中国 AI 发展：聚焦应用落地，推进技术适配与自主攻坚</b> .....	<b>10</b>
<b>4. 其他观点更新：AI+教育迎来国家级顶层设计</b> .....	<b>10</b>
<b>5. 投资建议</b> .....	<b>11</b>
<b>6. 风险提示</b> .....	<b>11</b>

## 图表目录

图 1: 我国 token 消耗量迅速提升 (单位: 万亿词元) .....	4
图 2: 2026 年中美通过 Openrouter 的词元消耗对比 (单位: 万亿词元) .....	4
表 1: 我国算力基础设施规划与目标达成情况.....	6

## 1. AI 商业价值爆发式增长：智能体驱动的 Token 经济革命，国产替代开启战略机遇期

2026 年，AI 产业商业价值迎来结构性裂变，核心驱动力为智能体 (Agent) 的全面商业化落地。AI 正从被动应答的对话式工具，加速演进为可自主调用工具、执行多步骤复杂任务的生产力引擎，直接带动 Token (词元) 消耗量呈指数级增长，进而重塑 AI 全产业的价值分配体系与底层增长逻辑。

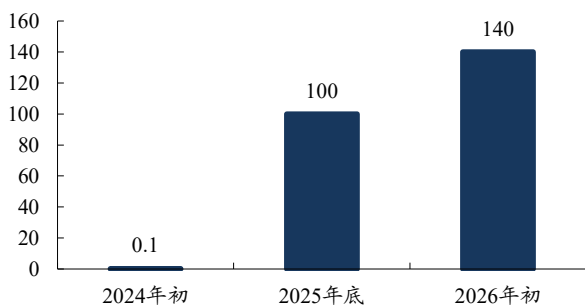
与此同时，DeepSeek V4 首次使用国产算力参与训练，成为本轮 AI 信创产业趋势中的关键事件。战略意义远超商业价值：中国 AI 产业链正在走向自主闭环，从依赖海外高端芯片与外部生态，转向芯片、算力、模型、框架、应用协同发展的自主可控体系。这一趋势一旦形成产业正反馈，将具备不可逆性。

### 1.1. Token 消耗爆发

智能体彻底重构了 AI 的使用范式。传统 Chatbot 单轮问答通常仅消耗数百 Token，而一个企业级智能体通常可完成从市场调研、报告撰写、PPT 生成到邮件分发的全流程，单次任务 Token 消耗量可达数十万至数百万级别，任务复杂度较传统对话场景大幅提升。

我国 Token 调用量已进入高速增长阶段。根据国家数据局披露数据，到 2026 年 3 月，我国日均词元 (Token) 调用量已超过 140 万亿，较 2024 年初的 1000 亿增长 1000 多倍，较 2025 年底的 100 万亿在三个月内增长超 40%；OpenRouter 最新周度数据看，中国大模型调用量在全球 API 生态中继续保持高活跃度。2026 年 4 月 27 日至 5 月 3 日，全球 AI 大模型总调用量为 23.9 万亿 Token，其中中国 AI 大模型周调用量升至 7.942 万亿 Token，环比增长 81.7%，时隔两周再次超过美国。国内大模型正在通过低价、免费版、长上下文和 Agent 能力，持续提升全球开发者生态中的调用份额。

图1：我国 token 消耗量迅速提升 (单位：万亿词元)



数据来源：新华社，东吴证券研究所

图2：2026 年中美通过 Openrouter 的词元消耗对比 (单位：万亿词元)



数据来源：每日经济新闻，东吴证券研究所

Token 消耗的持续爆发，正直接转化为 AI 企业营收规模的跃升，并深刻重塑 AI 市场竞争格局。头部 AI 企业收入规模持续高速增长，Token 已从单纯的技术计量单位，逐步演变为 AI 时代的核心生产资料与数字硬通货。此前全球科技巨头的产业竞争焦点在于模型参数量堆砌，而当下战略高地已转向 Token 生态体系构建；算力市场的需求重心也逐渐从训练主导转向推理主导。

## 1.2. DeepSeek V4 国产算力参与训练：AI 信创的战略拐点

DeepSeek V4 使用国产算力参与训练，是 AI 信创产业由政策驱动走向产业自证的重要拐点。

**第一，国产算力参与模型训练，国产 AI 闭环的起点。**从商业化角度看，国产算力训练短期并不具备优势。相比英伟达高端 GPU 体系，国产算力在训练效率、软件生态、集群稳定性、工程调优经验等方面仍存在差距，但选择国产算力的战略信号明确：国产算力开始进入高强度、大规模、复杂模型训练场景。这意味着国产算力产业链正在从可用向能用、好用、敢用演进，大模型厂商对国产算力的适配意愿和工程能力正在显著提升。

**第二，“十五五”规划明确提出实施“人工智能+”行动，强调推动核心技术自主可控，加快算力基础设施建设。**这是国家层面对 AI 国产化的顶层设计。从芯片到算力集群，从基础模型到行业应用，从政务场景到央企核心业务，全链条自主可控已成为 AI 产业发展的战略底线。DeepSeek V4 选择国产算力训练，本质上是产业端对上述顶层设计的回应和验证。

**第三，核心技术自主可控是基于当前 AI 产业竞争现实约束的战略。**习近平总书记在 2016 年网络安全和信息化工作座谈会上指出：“如果核心技术不是自己掌握的，那就像在别人的墙基上砌房子，再大再漂亮也可能经不起风雨，甚至会不堪一击。”在大模型成为新一代数字经济基础设施的背景下，底层算力若长期受制于人，将直接影响模型迭代、产业应用、数据安全和国家数字竞争力。DeepSeek V4 采用国产算力参与训练，正是对这一战略判断的产业实践。

**第四，竞争方的判断反向验证了国产算力替代的战略紧迫性。**英伟达 CEO 黄仁勋直言，若中国 AI 大模型率先跑在华为芯片上，对美国将是重大冲击。该表态从产业竞争维度说明，AI 算力国产替代不仅是中国自身的产业诉求，也是全球 AI 竞争格局中的关键变量。一旦中国大模型在国产芯片上形成稳定训练与推理闭环，海外高端芯片对中国 AI 产业发展的约束力将显著下降，国产算力产业链的战略价值也将迎来系统性重估。

因此，DeepSeek V4 事件的核心意义并不在于单一模型版本本身，而在于它标志着中国 AI 产业链正在走向国产模型+国产芯片的训推闭环，从局部替代走向体系替代，从技术追赶走向自主闭环。

## 2. AI 信创五大主线，国产替代战略价值显著抬升

全球科技封锁态势持续加深，国产替代的产业价值与战略意义显著抬升。2026 年以来，海外对华科技限制呈现高频化、体系化、长期化特征，AI 芯片、模型能力、技术交流、科研合作等环节均受到不同程度限制。上述事件一方面印证了中国 AI 企业正通过快速学习和产业化迭代缩小与海外领先水平的差距，另一方面也表明全球科技铁幕正在加速落地。

本轮海外对华技术封锁的核心目标，是尽可能延缓中国 AI 产业的技术进步节奏。虽然其无法阻挡中国人工智能产业的长期发展趋势，但短期内会对模型训练效率、先进算力供给、软件生态迭代形成扰动。在此背景下，AI 信创不再只是传统 IT 国产化的延伸，而是国家级科技竞争中的底层基础设施命题。

DeepSeek V4 使用国产算力参与训练，进一步强化了国产替代主线的产业确定性。过去市场对国产算力的理解，更多集中在政务信创、央国企采购、推理侧部署和低中高端训练场景；而此次事件表明，国产算力正在向大模型核心训练环节推进。即便当前国产算力在效率、成本和生态上仍有短板，但只要头部模型厂商愿意持续投入适配，国产算力产业链就有望在真实高强度场景中快速迭代，形成“模型牵引芯片、芯片反哺模型、应用放大需求”的正循环。

在此背景下，人工智能信创产业将扩展为五大主线：其一为 GPU 芯片，其二为 CPU 芯片，其三为昇腾产业链，其四为算力租赁，其五为信创大模型。国产算力替代趋势不可逆转，AI 信创产业链有望迎来从主题投资向产业兑现的阶段性切换。

### 2.1. 国产算力：基础设施扩容与国产替代同步推进

国内算力基础设施建设仍处于扩容阶段。工信部等六部门《算力基础设施高质量发展行动计划》提出，到 2025 年我国算力规模超过 300 EFLOPS，智能算力占比达到 35%，存储总量超过 1800 EB，先进存储容量占比达到 30%以上。而截至 2025H2，以上目标已基本完成，部分目标超额完成。

表1：我国算力基础设施规划与目标达成情况

指标	2025 年目标	公开可查最新实际值	完成度判断
算力规模	超过 300 EFLOPS	2025 年 6 月底，我国计算设备算力总规模 962 EFLOPS	超额完成
智能算力占比	达到 35%	2025 年 6 月底，智能算力占全国总算力比重达 81.9%； 2026 年 3 月底，我国智能算力规模达 1882 EFLOPS (FP16)	超额完成
存储总量	超过 1800 EB	2025 年 6 月底，存力规模超过 1680 EB	93.30%
先进存储容量占比	达到 30%以上	2025 年存力报告口径显示，全国外置闪存占比超过 28%	基本达成

数据来源：政务网，央广网，信通院，东吴证券研究所

**国产算力替代为国内算力基础设施行业发展的核心主线。**“十五五”规划提出实施“人工智能+”行动，加快建设现代化产业体系，推动人工智能与实体经济深度融合，同时强调核心技术自主可控与算力基础设施建设。这意味着算力国产化不再只是单点突破，而是将成为未来五年 AI 产业政策、资本开支和产业落地的共同主线。

**国产算力替代正在呈现推理侧先行+训练侧突破+生态侧协同的特征。**推理场景对稳定吞吐、成本效率和部署规模要求更高，是国产 AI 芯片率先放量的主要入口；**训练侧尤其是万亿参数级大模型，对芯片性能、互联能力、并行框架、算子生态和集群稳定性要求更高，国产替代难度更大。**DeepSeek V4 适配华为 Ascend 处理器，华为昇腾超节点已宣布支持 DeepSeek V4 系列，且华为芯片参与了 V4-Flash 部分训练环节，标志着国产算力开始从推理部署向训练链条渗透。若 DeepSeek V4 国产算力训练形成示范效应，则训练侧国产替代节奏存在进一步超预期可能。

### 2.1.1. GPU 芯片：国产算力替代的第一核心方向

**GPU 芯片是 AI 信创最核心、最直接的受益方向。**大模型训练和推理对高性能并行计算能力提出极高要求，GPU 及类 GPU 加速芯片是当前 AI 算力体系的关键底座。海外高端 GPU 供给受限后，国产 GPU 的重要性显著提升。

DeepSeek V4 选择国产算力参与训练，意味着国产 GPU 及 AI 加速芯片已从推理侧验证逐步走向训练侧验证。尽管当前国产 GPU 在软件生态、算子库、集群通信、稳定性和开发者工具链方面仍需持续完善，但只要进入真实大模型训练场景，产业迭代速度将明显加快。

**重点关注方向包括：寒武纪、海光信息、沐曦股份、摩尔线程等。**其中，寒武纪代表国产 AI 芯片核心标的；海光信息兼具高端 CPU 与 DCU 生态优势；沐曦股份、摩尔线程等国产 GPU 厂商有望在信创算力、智算中心和行业大模型训练推理场景中持续提升渗透率。

### 2.1.2. CPU 芯片：Agent 时代的底层调度核心

**本轮算力产业扩张周期中，除 GPU 赛道外，CPU 市场同样迎来显著需求增量。**Intel、AMD 相继启动 CPU 产品提价，全球服务器 CPU 市场已呈现显著供需错配格局。受先进制程产能挤占影响，Intel 与 AMD 2026 年产能配额已基本售罄。在供给端先进制程产能受限、生产成本持续上行的双重驱动下，两家巨头计划将全球产品报价上调 10%~15%。本轮面向全球云厂商与企业级客户的普涨，标志着服务器 CPU 市场已进入供给溢价阶段。

**AI 算力需求的超预期增长，叠加通用服务器更新周期的到来，持续推升 CPU 市场需求。** CPU 需求端的韧性主要来自两大核心支撑：其一，全球通用服务器已进入大规模存量更新窗口期，基座算力迭代需求为市场提供基本盘支撑；其二，大模型技术演进对 CPU 协同计算性能的要求持续超市场预期，单机 CPU 配置数量与性能规格同步提升，进一步放大市场供需缺口。在 AI 异构计算架构中，CPU 作为逻辑调度与 I/O 吞吐核心的产业地位愈发凸显。

**Agent 发展进一步重构 CPU 产业定位。** 与传统静态推理场景不同，Agent 系统需高频执行工具调用、沙盒隔离、多步规划与状态管理等操作，强化学习训练中的环境模拟与奖励计算环节，也高度依赖 CPU 的串行计算性能。伴随 Agent 与强化学习技术持续落地，AI 工作负载范式有望迎来重构，算力基础设施建设重心将从“GPU 堆叠算力”逐步转向“GPU+CPU 异构协同”。

**CPU 重点关注方向包括：禾盛新材、海光信息、中国长城、龙芯中科等。** 其中，海光信息在国产高端处理器与 AI 算力协同方面具备代表性；中国长城、龙芯中科受益于信创 CPU 生态持续推进；相关产业链公司有望在服务器、整机、政企信创和 AI 基础设施扩容中获得增量需求。

## 2.2. 昇腾链：国产 AI 算力生态的体系化主线

**昇腾产业链是 AI 信创中最具体系化特征的方向之一。** 与单一芯片替代不同，昇腾链条覆盖上游材料、连接器、PCB、服务器、整机、集成、算力平台和行业应用，是国产 AI 算力从芯片走向系统级落地的重要载体。

**DeepSeek V4 国产算力训练进一步提升市场对昇腾生态的关注度。** 若国产大模型在国产 AI 芯片上持续完成训练和推理闭环，昇腾链条的价值将不再局限于政策采购或行业信创项目，而是有望进入通用 AI 基础设施市场。国产算力生态的竞争，本质上不仅是芯片性能竞争，更是开发框架、算子库、通信网络、集群管理、模型适配和应用迁移效率的综合竞争。

**重点关注方向包括：云南锗业、华丰科技、航天电器、华正新材、神州数码等。** 上游材料、连接器、PCB 等环节受益于国产 AI 服务器和智算中心建设；神州数码等集成与渠道型厂商则有望受益于昇腾生态在政企、央企和行业客户中的推广。

## 2.3. 算力租赁：短期高弹性方向，商业模式迎来重估

**当前国内互联网大厂普遍采用自建/代建 AI 数据中心与算力租赁双轮推进的发展模式。** 核心原因在于，国内领先大模型的训练与推理仍高度依赖高端算力，而合规监管和海外供给约束使得大厂难以通过传统方式充分获取先进芯片。在此背景下，市场形成

显著供需缺口：高端算力需求刚性存在，但合规供给受限，第三方算力租赁厂商由此迎来发展机遇。

**Agent 规模化落地进一步拉动算力需求持续高增。**2026 年 Agent 智能体正逐步实现商业化落地，多模态大模型、编程大模型的 Token 消耗量较传统文本推理场景普遍增长 10 倍以上。伴随国内 Token 调用量持续陡峭增长，大模型厂商对稳定算力供给的需求持续强化。智谱 AI、DeepSeek 等大模型产品频繁出现服务压力，本质上反映出算力供给侧持续紧张。

**算力租赁商业模式正在迎来核心变革。**传统算力租赁模式下，厂商通常与互联网大厂签订 35 年期服务合同，在接到订单后约 5 个月内完成集群验收上线，随后按月收取租金。当前行业的关键变化在于，算力租赁厂商正从传统“裸算力出租”模式，逐步向参与大模型服务 Token 收入分成的模式演进。

在该模式下，算力租赁厂商不再只是基础设施出租方，而是大模型厂商 Token 收入体系中的重要参与方。大模型厂商的盈利能力不仅取决于模型性能，更需以稳定 Token 输出和算力支撑为前提。对于掌握稀缺算力供给的租赁厂商而言，提供算力服务本质上等同于为下游大模型厂商提供盈利基础。未来算力租赁厂商有望深度参与大模型厂商 Token 收入分成体系，估值逻辑也将从传统“租金收益型”切换为“Token 收入分成型”。

**重点关注方向包括：宏景科技、协创数据、盈峰环境、利通电子等。**行业壁垒主要体现在三方面：第一，稳定的芯片获取能力与交付效率；第二，卓越的集群组网与性能调优能力；第三，雄厚的资金实力与高效的资金周转能力。伴随行业参与者数量减少、订单向头部集中，算力租赁赛道将呈现显著马太效应。

## 2.4. 信创大模型：党政、地方政府与央国企需求尚未充分定价

在 AI 信创产业发展逻辑下，当前市场尚未充分重视的核心赛道为党政机关、地方政府与央国企的信创大模型需求。大模型作为数字经济时代的核心基础设施，其自主可控具备重要战略意义。

相关政策明确提出，要提高政府治理数智化水平，深化数智技术全流程应用，发展泛在可及、智慧便捷、公平普惠的数智化政务服务；安全稳妥有序推进政务领域人工智能大模型部署应用，探索构建精准识别需求、主动规划服务、全程智能办理的服务新模式；深化人工智能赋能安全治理，提升感知预警、指挥决策、精准管理和即时响应能力。

**政务、央国企和重点行业大模型对安全性、可控性、私有化部署、数据闭环和国产算力适配提出更高要求。**这类市场不会简单采用海外闭源模型或依赖不可控算力体系，而更倾向于选择具备国产算力适配能力、本地化交付能力和行业 Knowhow 的大模型厂商。

在政务与央国企信创大模型市场，科大讯飞星火大模型具备显著竞争优势。星火大模型坚持国产算力路线，结合自研大模型并行训练框架，持续提升模型训练性能。科大讯飞在教育、政务、央国企等场景长期积累数据、渠道和项目交付能力，有望在 AI 信创大模型市场持续受益。

同时，智谱、MiniMax 等国内大模型企业也值得重点关注。进入 Agent 时代后，大模型在全新商业模式下正成为 AI 产业最重要的基础设施。随着国产算力训练和推理能力不断验证，信创大模型有望从“可选项”逐步变为政企客户数字化转型的“必选项”。

### 3. 中国 AI 发展：聚焦应用落地，推进技术适配与自主攻坚

当前全球人工智能大模型产业处于技术快速迭代、产品密集发布的发展阶段，中美两国在 AI 技术研发与应用落地的核心路径上，已形成鲜明差异化特征。美国市场更聚焦前沿技术突破与生产力核心瓶颈攻坚，中国市场更侧重将成熟 AI 技术转化为可规模化落地的实用型生产力工具，同时持续补齐基础技术短板。

中国 AI 产业的核心发力点，在于将现有前沿 AI 技术转化为可落地、可规模化应用的生产力工具，在消费内容生成、行业场景数字化落地等领域实现快速推进，同时持续推进国产算力适配与基础技术自主创新。

DeepSeek V4 国产算力训练进一步说明，中国 AI 产业的追赶路径正在发生变化。过去中国大模型更多依赖海外高端芯片和开源技术生态进行快速迭代，当前则正在走向以国产算力为底座、以本土数据和应用场景为牵引、以行业落地为商业化出口的自主发展路径。虽然国内通用大模型整体仍处于追赶阶段，但在工程效率、应用适配、成本控制、场景落地和国产化闭环方面，中国 AI 产业具备独特优势。

未来国内 AI 产业有望形成两条并行主线：一是继续追赶全球前沿模型能力，缩小通用智能水平差距；二是围绕国产算力、信创大模型、行业 Agent 和政企应用构建自主可控生态。后者将是 AI 信创的核心投资方向。

### 4. 其他观点更新：AI+教育迎来国家级顶层设计

教育部、国家发改委、工信部、科技部、国家数据局五部门联合发布《“人工智能+教育”行动计划》，并于 2026 年 4 月 10 日正式对外公布。该政策并非单纯试点通知，而是“十五五”开局之年，国家为教育数字化转型制定的顶层路线图，核心目标是推动人工智能全面融入教育全要素、全过程、全场景。

AI+教育是信创大模型最具确定性的落地场景之一。教育场景具备高频刚需、数据持续沉淀、预算来源明确、政策支持力度强等特点，且对数据安全、内容安全、模型可

控性要求较高，因此更适合国产大模型和国产算力体系落地。

科大讯飞智慧教育业务已在大量学校实现深度应用，服务师生规模领先，具备数据、渠道、产品和交付能力优势。伴随国家财政教育建设资金投入方向从传统信息化建设逐步向 AI 教育迁移，AI 学习机、教育大模型、区域教育智能体、智慧课堂等产品有望持续放量。教育场景也将成为检验国产大模型行业落地能力的重要样板。

## 5. 投资建议

综合来看，国产算力替代趋势不可逆转。DeepSeek V4 采用国产算力训练，标志着 AI 信创从政策主题走向产业实践，从推理侧替代走向训练侧突破，从局部国产化走向全链条自主可控。建议重点关注 AI 信创五大方向：

**第一，GPU 芯片。**建议关注寒武纪、海光信息、沐曦股份、摩尔线程。国产 GPU 及 AI 加速芯片是大模型训练和推理的核心底座，DeepSeek V4 事件有望加速市场对国产训练算力的认知重估。

**第二，CPU 芯片。**建议关注禾盛新材、海光信息、中国长城、龙芯中科。Agent 和强化学习的发展提升了 CPU 在 AI 异构架构中的战略地位，高性能 CPU 将成为 AI 基础设施的重要瓶颈环节。

**第三，昇腾链。**建议关注云南锗业、华丰科技、航天电器、华正新材、神州数码。昇腾生态代表国产 AI 算力体系化落地的重要方向，受益于国产 AI 服务器、智算中心和政企信创需求扩张。

**第四，算力租赁。**建议关注宏景科技、协创数据、盈峰环境、利通电子。短期看，算力供需紧张推动租赁价格上行；中期看，Token 收入分成模式有望带来商业模式重估。

**第五，信创大模型。**建议关注科大讯飞、智谱、MiniMax。政务、央企、教育、金融、能源等行业对自主可控大模型需求持续提升，国产算力适配能力将成为核心竞争壁垒。

## 6. 风险提示

**政策落地不及预期：**若财政预算、招采节奏放缓，AI 信创需求释放或低于预期。

**技术迭代不及预期：**若国产芯片、软件生态及模型适配进展较慢，替代节奏或受影响。

**算力供给波动风险：**若高端芯片获取、集群交付或运维稳定性不达预期，项目落地或延迟。

**商业化兑现风险：**若 Token 调用增长难以转化为收入，算力租赁和大模型盈利或承压。

**市场竞争加剧风险：**若模型厂商、芯片厂商及云服务商价格竞争加剧，行业利润率或下行。

## 免责声明

东吴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。

本研究报告仅供东吴证券股份有限公司（以下简称“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，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。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。

在法律许可的情况下，东吴证券及其所属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投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中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。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

## 东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

- 买入：预期未来 6 个月个股涨跌幅相对基准在 15%以上；
- 增持：预期未来 6 个月个股涨跌幅相对基准介于 5%与 15%之间；
- 中性：预期未来 6 个月个股涨跌幅相对基准介于-5%与 5%之间；
- 减持：预期未来 6 个月个股涨跌幅相对基准介于-15%与-5%之间；
- 卖出：预期未来 6 个月个股涨跌幅相对基准在-15%以下。

行业投资评级：

- 增持：预期未来 6 个月内，行业指数相对强于基准 5%以上；
- 中性：预期未来 6 个月内，行业指数相对基准-5%与 5%；
- 减持：预期未来 6 个月内，行业指数相对弱于基准 5%以上。

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应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所  
苏州工业园区星阳街 5 号  
邮政编码：215021  
传真：（0512）62938527  
公司网址：<http://www.dwzq.com.cn>