



Research and
Development Center

Deepseek R1 是 AGI 的里程碑，中长期利好 算力硬件

2025 年 2 月 4 日

证券研究报告

行业研究

行业专题研究（普通）

电子

投资评级 看好

上次评级 看好

莫文宇 电子行业首席分析师

执业编号：S1500522090001

邮箱：mowenyu@cindasc.com

信达证券股份有限公司

CINDA SECURITIES CO., LTD

北京市西城区宣武门西大街甲127号金隅大厦
B座

邮编：100031

Deepseek R1 是 AGI 的里程碑，中长期利好算力硬件

2025 年 2 月 4 日

本期内容提要：

- **DeepSeek 树立里程碑，挖掘算法创新的价值潜力。** Deepseek V3 仍是基于 Transformer 架构的模型，是一个强大的专家混合（MoE）语言模型，总共有 671B 个参数，每个令牌激活了 37B 参数。DeepSeek-V3 主要贡献包括：架构层面创新的负载均衡策略和训练目标，预训练层面大幅提高训练效率，后训练层面 DeepSeek-R1 的知识提炼等。DeepSeek 团队仅以 2.664M H800 GPU 小时的经济成本，在 14.8T tokens 上完成了 DeepSeek V3 的预训练。**Deepseek R1 是基于 Deepseek V3 的架构上的集大成之作，性能对标 OpenAI-o1。** DeepSeek-R1-Zero 是一种通过大规模强化学习（RL）训练的模型，没有监督微调（SFT）作为初步步骤，展示了卓越的推理能力。通过强化学习，DeepSeek-R1-Zero 自然而然地出现了许多强大而有趣的推理行为。其中，**DeepSeek-R1-Zero 展示了自我验证、反射和生成长 CoT 等功能，这标志着研究界的重要里程碑。**值得注意的是，这是第一个验证的开放研究，可以纯粹通过 RL 来激励的 LLMs 推理能力，而无需 SFT。这一突破有望为该领域的未来发展铺平道路。此外，DeepSeek 团队开源了 DeepSeek-R1-Zero、DeepSeek-R1 以及基于 Qwen 和 Llama 从 DeepSeek-R1 中提炼出来的六个密集模型（1.5B、7B、8B、14B、32B、70B），在部分测试上性能超越 OpenAI o1-mini。
- **OpenAI 掀起反攻号角，o3-mini 和 Deep Research 相继发布。**为了应对 DeepSeek 的挑战，OpenAI 发布了 O3-mini，理解物理世界能力、编程能力大幅提升。o3-mini 在物理模拟高难度挑战中展现出惊人的实力，在模拟小球的转动时，o3-mini 具备更强的物理推理能力，而 DeepSeek R1 则出现了反重力现象。此外，o3-mini 能生成四维空间内小球的弹射程序，展现出很高的潜力。此外，**OpenAI 还发布了 Deep Research，这是 OpenAI 最新发布代理——**仅需一个提示，ChatGPT 将查找、分析和综合数百个在线资源，以创建研究分析师级别的综合报告。它由即将推出的 OpenAI o3 模型版本提供支持，该版本针对 Web 浏览和数据分析进行了优化，它利用推理来搜索、解释和分析互联网上的大量文本、图像和 PDF，并根据遇到的信息根据需要进行调整。
- **AI 未来尚有星辰大海有待探索，AGI 商业化奇点临近。**谷歌 DeepMind 团队曾将 AI 划分为 6 大发展阶段，从狭义的层面看，在专业领域已经有 AI 模型能达到超人级别。如 AlphaFold、AlphaZero、StockFish 等 AI 模型可以在特定的领域完全超越人类，但是从通用人工智能的层面看，AI 的发展还处于较低水平，如 ChatGPT 仅可以划分为 Level 1-Emerging 级别。**FIGURE 是人形机器人领域的领导者之一，在其视角中，我们也能窥探 AI 的历史机遇。在如今全球 GDP 中，50% 的劳动力均为人类劳动，大约 42T 美金空间，仅此一项，AI 就有广阔的前景。**
- **投资建议：算法创新和算力投入相辅相成，建议关注 AI 产业链标的。** AI 领域的发展速度超过传统制造业，2023 年来，Scaling Laws 开始发力，全球领域内开启了算力资源的“军备竞赛”。而在地缘政治等因素的激励下，算法层面的追赶终于由算力相对匮乏的中国团队唱响号角。我们认为，在通向 AGI 的道路上，算法创新和算力资源都是不可或缺的。早在 2024 年，我们就明确了推理市场要想实现繁荣必须降低推理成本的结论。如今，豆包、DeepSeek 研究团队的相关工作成功

地大幅度降低了推理成本，推进了 AI 的产业化。DeepSeek 的成功，是开源模型相对闭源模型的一次飞跃。诚然，在十分有限的时间区间内，北美的科技大厂或把有限的精力放在算法层面，以充分挖掘自身算力资源的潜力。但是，这并不意味着对算力的投入将会停止。相反，AI 的产业化有望在中长期加速算力的投入，避免了以往的 AI 热潮最终因无法落地而走向失败的困局。从推理市场看：推理有望快速放量，文生文、文生视频、图生视频的各类功能或快速迭代，并使得 AI 真正的理解物理世界，自动驾驶、人形机器人、AI 在各行各业掀起产业革命的奇点临近。从训练市场看：一方面，训练模型的前沿探索仍需大量算力投入，且推理市场的放量，有望衍生新的模型探索方向。另一方面，如世界模型等方向的研究进展有望加速。但是，我们也看到大模型厂商的淘汰也将加速，无法超越开源模型的闭源大模型厂商或快速出清。扎克伯格曾在 Meta 法说会上表示，“随着时间的推移，就像每个企业都有一个网站、一个社交形象和一个电子邮件地址一样，在未来，每个企业也将拥有一个客户可以与之交互的 AI 代理。我们的目标是让每个小企业，最终每个企业，都可以轻松地将其所有内容和目录提取到一个 AI 代理中，从而推动销售并节省资金。”这一时点正在逐步临近，从历史看，当计算机从实验室走向千家万户，并没有导致相关的企业没落，而是孕育了许多伟大的企业。目前看，AI 也有望经历这一过程，建议关注产业链相关标的。

- **相关个股：**【AI 云侧】工业富联/沪电股份/生益科技/深南电路/胜宏科技/寒武纪/海光信息；【AI 端侧】蓝思科技/领益智造/鹏鼎控股/东山精密/乐鑫科技/瑞芯微/恒玄科技/全志科技/兆易创新/晶晨股份等。
- **风险因素：**宏观经济下行风险；下游需求不及预期风险；中美贸易摩擦加剧风险。

目录

DeepSeek 算法创新众多，性能表现较好.....	5
OpenAI 掀起反攻号角，o3-mini 和 Deep Research 相继发布.....	9
AI 未来尚有星辰大海有待探索，AGI 商业化奇点临近.....	13
风险因素.....	15

表目录

表 1: 建议关注个股.....	15
------------------	----

图目录

图 1: Deepseek V3 基本架构.....	6
图 2: DeepSeek-R1-Zero 的“啊哈时刻”.....	7
图 3: deepseek 性能对标 OpenAI o1.....	8
图 4: deepseek 蒸馏小模型在部分测试上性能超越 OpenAI o1-mini.....	8
图 5: o3-mini 具备更强的反重力推理能力.....	9
图 6: o3-mini 生成了小球在四维空间弹射的代码.....	9
图 7: o3-mini 理解物理世界的 demo.....	9
图 8: o3-mini 仅用 8 秒时间就可克隆一个应用.....	10
图 9: o3-mini 生成的贪吃蛇游戏.....	10
图 10: o3-mini 生成的射击游戏.....	10
图 11: o3-mini 生成的漂浮城市.....	10
图 12: o3-mini 定价对比.....	11
图 13: deep research 在人类终极考试中的成绩.....	12
图 14: deep research 专家级别任务通过率.....	12
图 15: 谷歌提出的 AGI 等级.....	13
图 16: FIGURE 看到的长期机会.....	14

DeepSeek 树立里程碑，挖掘算法创新的价值潜力

Deepseek V3 仍是基于 Transformer 架构的模型，是一个强大的专家混合（MoE）语言模型，总共有 671B 个参数，每个令牌激活了 37B 参数。实现高效的推理和具有成本效益的训练，DeepSeek-V3 采用了多头潜在注意力（MLA）和 DeepSeekMoE 架构，此外，DeepSeek-V3 开创了一种用于负载均衡的辅助无损策略，并设定了多标记预测训练目标以获得更强的性能。尽管性能出色，但 DeepSeek-V3 只需要 2.788M H800 GPU 小时即可进行完整训练。总结看，DeepSeek-V3 主要贡献包括：

架构层面：创新的负载均衡策略和训练目标。

- 除了 DeepSeek-V2 的高效架构之外，DeepSeek V3 还开创了一种用于负载均衡的辅助无损策略，该策略可以最大限度地减少因鼓励负载而引起的性能下降。
- 团队研究了多标记预测（MTP）目标，并证明对模型性能有益，还可用于推理加速的推测解码。

预训练：迈向终极训练效率。

- Deepseek 设计了 FP8 混合训练精度训练框架，并且首次在超大规模模型上验证了 FP8 训练的可行性和有效性。
- 通过算法、框架和硬件的协同设计，DeepSeek 克服了跨节点 MoE 训练中的通信瓶颈，实现了近乎全计算通信折叠，显著提高了 DeepSeek 的训练效率并且降低训练成本，使我们能够在不增加开销的情况下进一步扩大模型大小。
- DeepSeek 仅以 2.664M H800 GPU 小时的经济成本，在 14.8T tokens 上完成了 DeepSeek V3 的预训练，生成了当时最强的对外发布的开源基础模型。

后训练：DeepSeek-R1 的知识提炼。

- DeepSeek 引入了一种创新的方法，将 CoT 的推理能力提炼出来，从 DeepSeek-R1 模型中提取推理能力并应用于 DeepSeek-V3，提升其推理性能。

图 1: Deepseek V3 基本架构

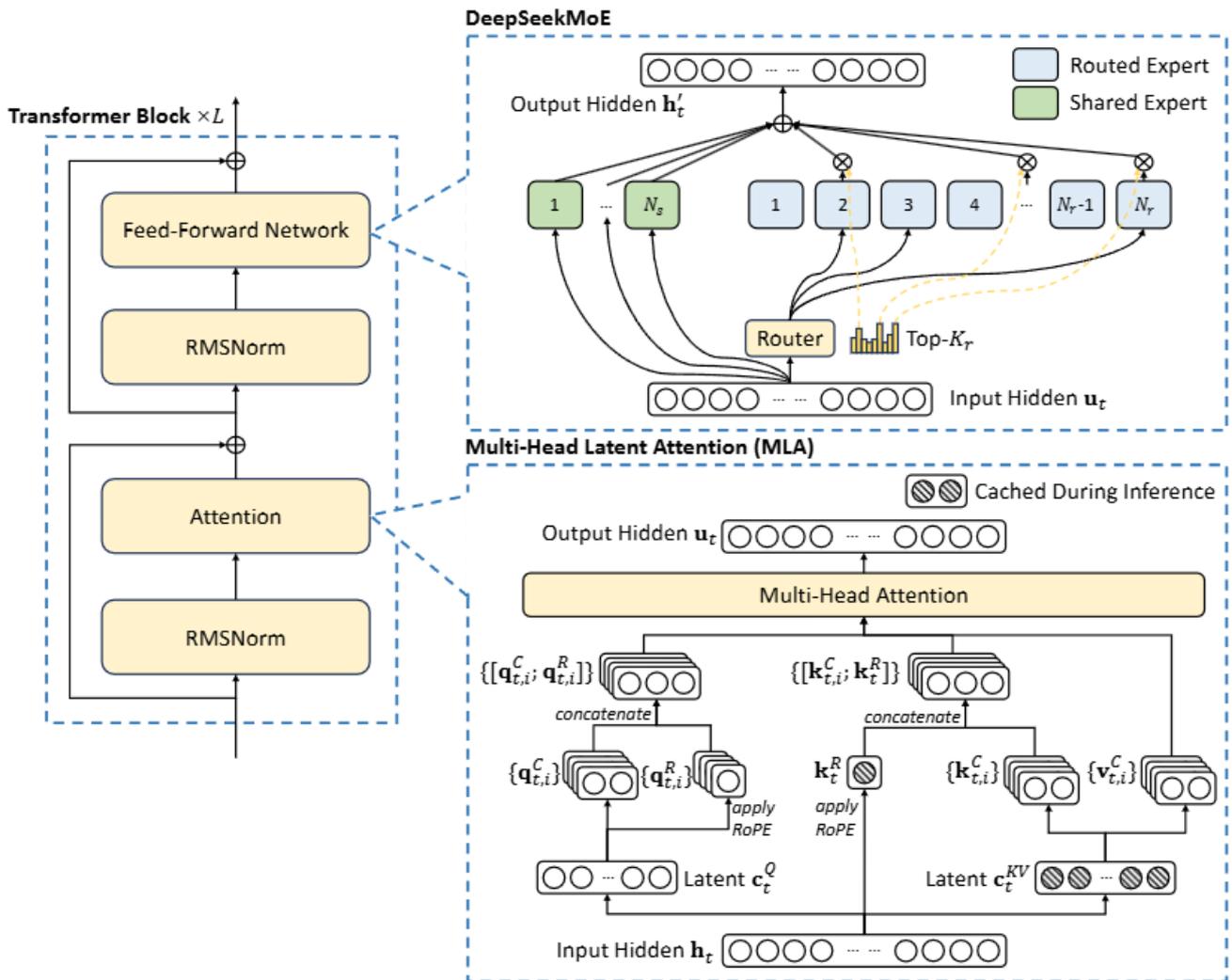


Figure 2 | Illustration of the basic architecture of DeepSeek-V3. Following DeepSeek-V2, we adopt MLA and DeepSeekMoE for efficient inference and economical training.

资料来源: DeepSeek-AI 《DeepSeek-V3 Technical Report》, 信达证券研发中心

Deepseek R1 是基于 Deepseek V3 的架构上的集大成之作。 DeepSeek-R1-Zero 是一种通过大规模强化学习 (RL) 训练的模型, 没有监督微调 (SFT) 作为初步步骤, 展示了卓越的推理能力。通过强化学习, DeepSeek-R1-Zero 自然而然地出现了许多强大而有趣的推理行为。但是, 它遇到了可读性差和语言混合等挑战。为了解决这些问题并进一步提高推理性能, DeepSeek 团队引入了 DeepSeek-R1, 它在 RL 之前结合了多阶段训练和冷启动数据。DeepSeekR1 在推理任务上实现了与 OpenAI-o1-1217 相当的性能。为了支持研究社区, DeepSeek 团队开源了 DeepSeek-R1-Zero、DeepSeek-R1 以及基于 Qwen 和 Llama 从 DeepSeek-R1 中提炼出来的六个密集模型 (1.5B、7B、8B、14B、32B、70B)。

后训练: 在基础模型上进行大规模强化学习。

- DeepSeek 直接将 RL 用于基础模型, 而无需以来监督微调作为初步步骤。这种方法允许模型探索解决复杂问题的思维链 CoT, 从而开发 DeepSeek-R1-Zero。DeepSeek-R1-Zero 展示了自我验证、反射和生成长 CoT 等功能, 这标志着研究界的重要里程碑。

值得注意的是，这是第一个验证的开放研究，可以纯粹通过 RL 来激励的 LLMs 推理能力，而无需 SFT。这一突破有望为该领域的未来发展铺平道路。

- DeepSeek 介绍了开发 DeepSeek-R1 的管道。该管道包含两个 RL 阶段，旨在发现改进的推理模式并与人类偏好保持一致，以及两个 SFT 阶段，作为模型推理和非推理能力的种子。

蒸馏：较小的模型也可以很强大。

- DeepSeek 证明，较大模型的推理模式可以提炼成较小的模型，与通过 RL 在小型模型上发现的推理模式相比，性能更好。开源 DeepSeek-R1 及其 API 有望使研究社区在未来能够提炼出更好的更小模型。
- 使用 DeepSeek-R1 生成的推理数据，DeepSeek 微调了研究界广泛使用的几个密集模型。评估结果表明，蒸馏的较小密集模型在基准上表现非常出色。DeepSeekR1-Distill-Qwen-7B 在 AIME 2024 上取得了 55.5% 的成绩，超过了 QwQ-32B-Preview。此外，DeepSeek-R1-Distill-Qwen-32B 在 AIME 2024 上的得分为 72.6%，在 MATH-500 上为 94.3%，在 LiveCodeBench 上为 57.2%。这些结果明显优于以前的开源模型，可与 o1-mini 相媲美。DeepSeek 开源了基于 Qwen2.5 和 Llama3 系列的 1.5B、7B、8B、14B、32B 和 70B 检查点给社区。

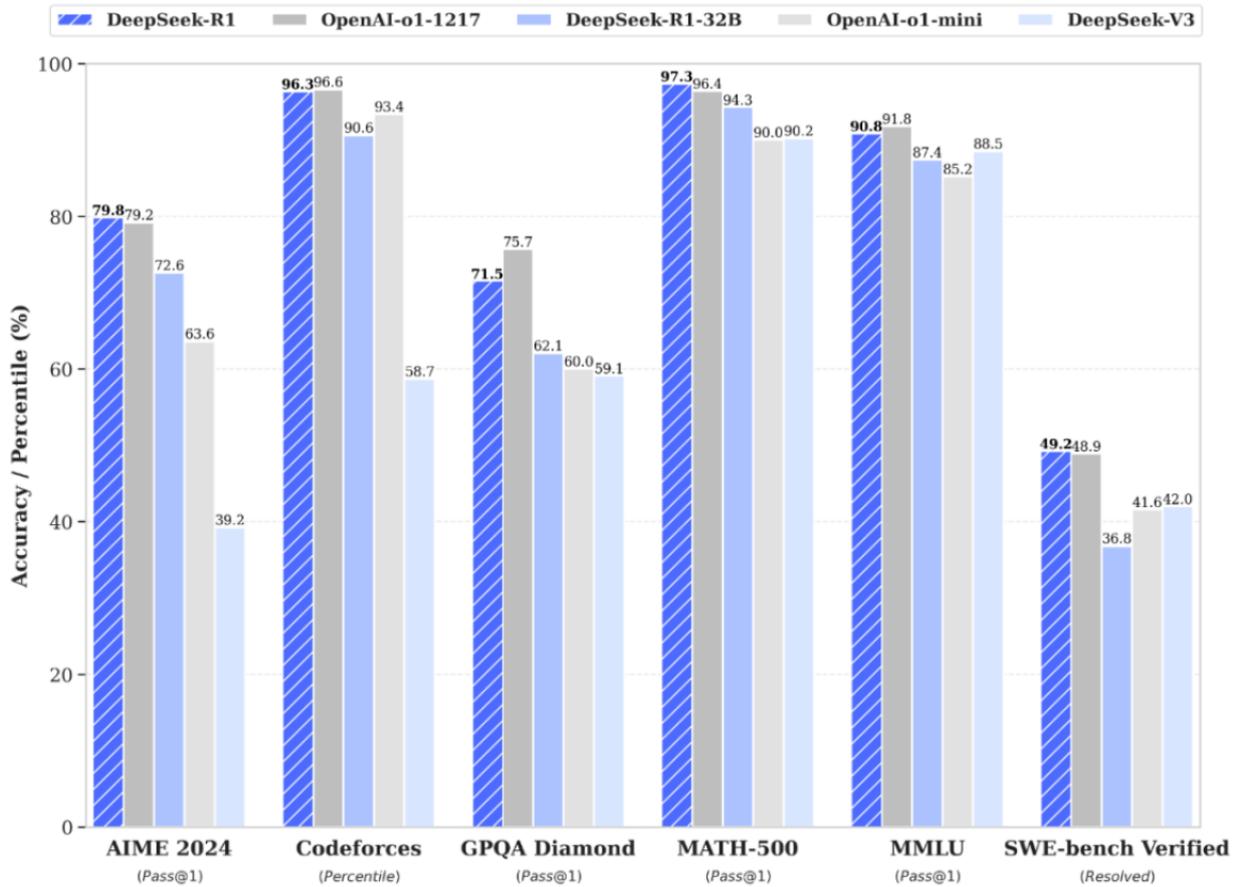
图 2: DeepSeek-R1-Zero 的“啊哈时刻”

Question: If $a > 1$, then the sum of the real solutions of $\sqrt{a - \sqrt{a+x}} = x$ is equal to

Response: <think>
 To solve the equation $\sqrt{a - \sqrt{a+x}} = x$, let's start by squaring both ...
 $(\sqrt{a - \sqrt{a+x}})^2 = x^2 \implies a - \sqrt{a+x} = x^2$.
 Rearrange to isolate the inner square root term:
 $(a - x^2)^2 = a + x \implies a^2 - 2ax^2 + (x^2)^2 = a + x \implies x^4 - 2ax^2 - x + (a^2 - a) = 0$
 ...
Wait, wait. Wait. That's an aha moment I can flag here.
 Let's reevaluate this step-by-step to identify if the correct sum can be ...
 We started with the equation:
 $\sqrt{a - \sqrt{a+x}} = x$
 First, let's square both sides:
 $a - \sqrt{a+x} = x^2 \implies \sqrt{a+x} = a - x^2$
 Next, I could square both sides again, treating the equation: ...
 ...

Table 3 | An interesting “aha moment” of an intermediate version of DeepSeek-R1-Zero. The model learns to rethink using an anthropomorphic tone. This is also an aha moment for us, allowing us to witness the power and beauty of reinforcement learning.

资料来源: EpochAI, 信达证券研发中心

图 3: deepseek 性能对标 OpenAI o1


资料来源: EpochAI, 信达证券研发中心

图 4: deepseek 蒸馏小模型在部分测试上性能超越 OpenAI o1-mini

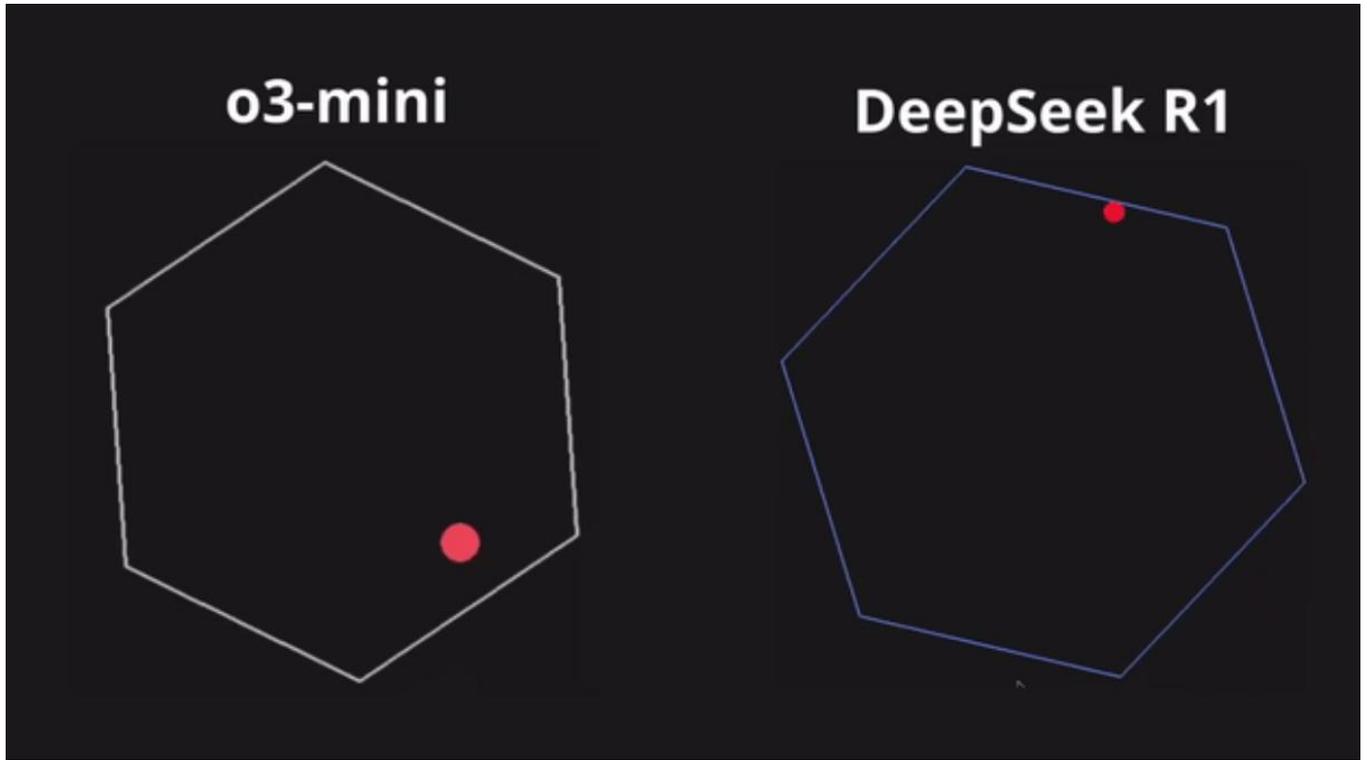
	AIME 2024 pass@1	AIME 2024 cons@64	MATH-500 pass@1	GPQA Diamond pass@1	LiveCodeBench pass@1	CodeForces rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759.0
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717.0
o1-mini	63.6	80.0	90.0	60.0	53.8	1820.0
QwQ-32B	44.0	60.0	90.6	54.5	41.9	1316.0
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954.0
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189.0
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481.0
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1691.0
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205.0
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1633.0

资料来源: EpochAI, 信达证券研发中心

OpenAI 掀起反攻号角，o3-mini 和 Deep Research 相继发布

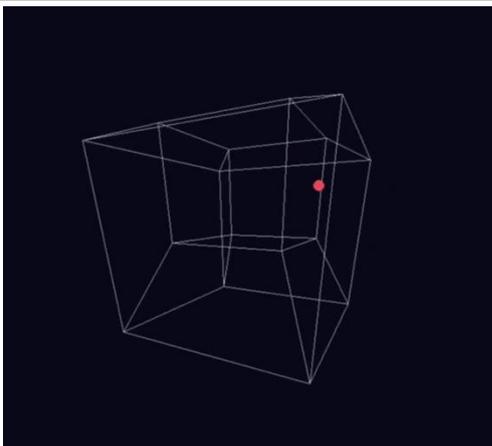
O3-mini 发布，理解物理世界能力、编程能力大幅提升。o3-mini 在物理模拟高难度挑战中展现出惊人的实力，在模拟小球的转动时，o3-mini 具备更强的物理推理能力，而 DeepSeek R1 则出现了反重力现象。此外，o3-mini 能生成四维空间内小球的弹射程序，展现出很高的潜力。

图 5: o3-mini 具备更强的反重力推理能力



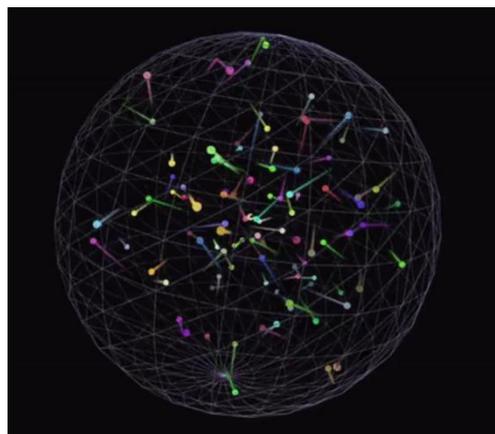
资料来源：机器之心，信达证券研发中心

图 6: o3-mini 生成了小球在四维空间弹射的代码



资料来源：EpochAI，信达证券研发中心

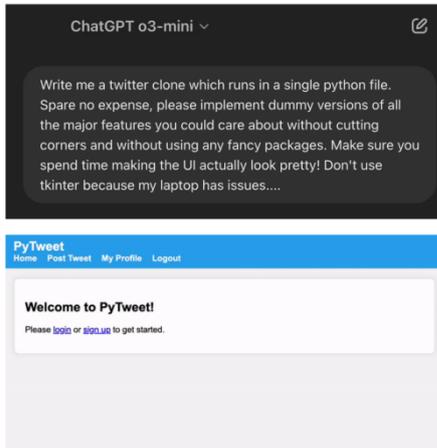
图 7: o3-mini 理解物理世界的 demo



资料来源：EpochAI，信达证券研发中心

编程能力方面，用户仅需一个提示，o3-mini 就能达到克隆应用、生成游戏的效果。从用户的评测看，用户要求 o3-mini 用单个 Python 文件写一个 Twitter 克隆应用，o3-mini 实现整个过程只用了 8 秒时间。此外，o3-mini 还可以用单个提示生成贪吃蛇、射击等游戏，编程能力让人惊叹。

图 8: o3-mini 仅用 8 秒时间就可克隆一个应用



资料来源: 机器之心, 信达证券研发中心

图 9: o3-mini 生成的贪吃蛇游戏



资料来源: 机器之心, 信达证券研发中心

图 10: o3-mini 生成的射击游戏



资料来源: EpochAI, 信达证券研发中心

图 11: o3-mini 生成的漂浮城市



资料来源: EpochAI, 信达证券研发中心 (Prompt: build me a massive amazing epic a floating city)

定价方面，o3-mini 每百万 tokens 的输入价格为 0.55 美元（缓存命中）/1.1 美元（缓存未命中），输出价格为 4.4 美元，相对 DeepSeek R1 仍然较贵。

图 12: o3-mini 定价对比

模型API定价 (每百万tokens)		
模型	输入价格	输出价格
DeepSeek-R1	0.14美元 (缓存命中)	2.19美元
	0.55美元 (缓存未命中)	
OpenAI o3-mini	0.55美元 (缓存命中)	4.40美元
	1.10美元 (缓存未命中)	
OpenAI o1-mini	0.55美元 (缓存命中)	4.40美元
	1.10美元 (缓存未命中)	
OpenAI o1	7.50美元 (缓存命中)	60.00美元
	15.00美元 (缓存未命中)	
DeepSeek-V3	0.014美元 (缓存命中)	0.28美元
	0.14美元 (缓存未命中)	
OpenAI GPT-4o-mini	0.075美元 (缓存命中)	0.60美元
	0.15美元 (缓存未命中)	
OpenAI GPT-4o	1.25美元 (缓存命中)	10.00美元
	2.50美元 (缓存未命中)	

资料来源: 智东西, 信达证券研发中心

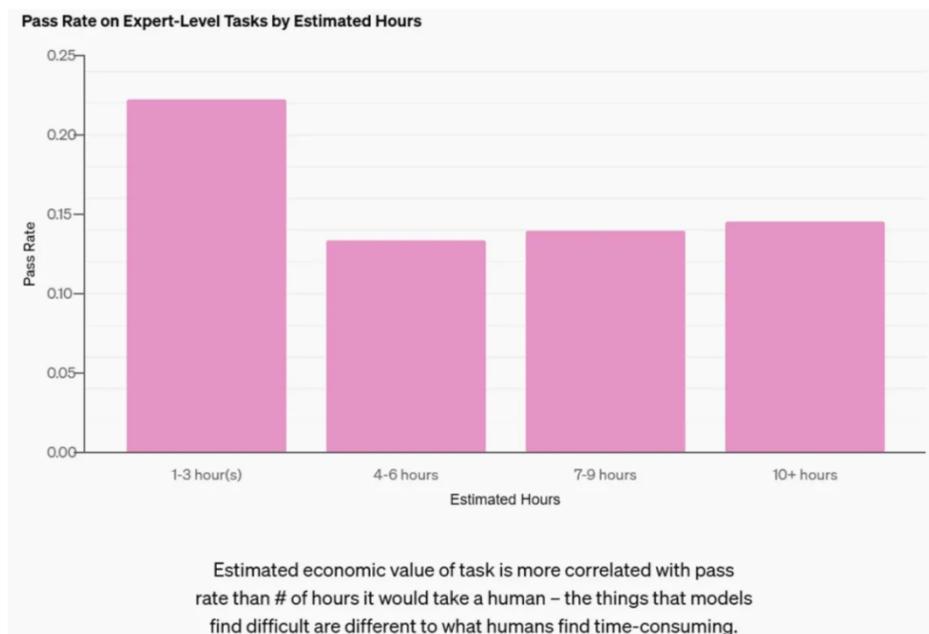
Deep Research 是 OpenAI 新发布代理——仅需一个提示, ChatGPT 将查找、分析和综合数百个在线资源, 以创建研究分析师级别的综合报告。它由即将推出的 OpenAI o3 模型版本提供支持, 该版本针对 Web 浏览和数据分析进行了优化, 它利用推理来搜索、解释和分析互联网上的大量文本、图像和 PDF, 并根据遇到的信息根据需要进行调整。

图 13: deep research 在人类终极考试中的成绩

Model	Accuracy (%)
GPT-4o	3.3
Grok-2	3.8
Claude 3.5 Sonnet	4.3
Gemini Thinking	6.2
OpenAI o1	9.1
DeepSeek-R1*	9.4
OpenAI o3-mini (medium)*	10.5
OpenAI o3-mini (high)*	13.0
OpenAI deep research**	26.6

* Model is not multi-modal, evaluated on text-only subset.
 **with browsing + python tools

资料来源: OpenAI 官网, 信达证券研发中心

图 14: deep research 专家级别任务通过率


资料来源: OpenAI 官网, 信达证券研发中心

AI 未来尚有星辰大海有待探索，AGI 商业化奇点临近

谷歌 DeepMind 团队曾将 AI 划分为 6 大发展阶段，每个阶段都有对应的深度（性能）和广度（通用性）指标。

- 第一阶段为“涌现级”（Emerging），技能相当于初出茅庐的人类。
- 第二阶段可理解为“合格级”（Competent），可以达到熟练成年人 50% 的水平。
- 第三阶段为“专家级”（Expert），可达到正常成年人 90% 的水平。
- 第四阶段为“大师级”（Virtuoso），可达到正常人类 99% 的水平。
- 第五阶段为“超人级”（Superhuman），在超越人类。

从狭义的层面看，在专业领域已经有 AI 模型能达到超人级别。如 AlphaFold、AlphaZero、StockFish 等 AI 模型可以在特定的领域完全超越人类，但是从通用人工智能的层面看，AI 的发展还处于较低水平，如 ChatGPT 仅可以划分为 Level 1-Emerging 级别。

图 15: 谷歌提出的 AGI 等级

Levels of AGI

Performance (rows) x Generality (columns)	Narrow <i>clearly scoped task or set of tasks</i>	General <i>wide range of non-physical tasks, including metacognitive abilities like learning new skills</i>
Level 0: No AI	Narrow Non-AI calculator software; compiler	General Non-AI human-in-the-loop computing, e.g., Amazon Mechanical Turk
Level 1: Emerging <i>equal to or somewhat better than an unskilled human</i>	Emerging Narrow AI GOFAI ⁴ ; simple rule-based systems, e.g., SHRDLU (Winograd, 1971)	Emerging AGI ChatGPT (OpenAI, 2023), Bard (Anil et al., 2023), Llama 2 (Touvron et al., 2023)
Level 2: Competent <i>at least 50th percentile of skilled adults</i>	Competent Narrow AI toxicity detectors such as Jigsaw (Das et al., 2022); Smart Speakers such as Siri (Apple), Alexa (Amazon), or Google Assistant (Google); VQA systems such as PaLI (Chen et al., 2023); Watson (IBM); SOTA LLMs for a subset of tasks (e.g., short essay writing, simple coding)	Competent AGI not yet achieved
Level 3: Expert <i>at least 90th percentile of skilled adults</i>	Expert Narrow AI spelling & grammar checkers such as Grammarly (Grammarly, 2023); generative image models such as Imagen (Saharia et al., 2022) or Dall-E 2 (Ramesh et al., 2022)	Expert AGI not yet achieved
Level 4: Virtuoso <i>at least 99th percentile of skilled adults</i>	Virtuoso Narrow AI Deep Blue (Campbell et al., 2002), AlphaGo (Silver et al., 2016, 2017)	Virtuoso AGI not yet achieved
Level 5: Superhuman <i>outperforms 100% of humans</i>	Superhuman Narrow AI AlphaFold (Jumper et al., 2021; Varadi et al., 2021), AlphaZero (Silver et al., 2018), StockFish (Stockfish, 2023)	Artificial Superintelligence (ASI) not yet achieved

量子位

资料来源：量子位，信达证券研发中心

FIGURE 是人形机器人领域的领导者之一，在其视角中，我们也能窥探 AI 的历史机遇。在如今全球 GDP 中，50%的劳动力均为人类劳动，大约 42T 美金空间，仅此一项，AI 就有广阔的前景。

图 16: FIGURE 看到的长期机会



资料来源: FIGURE 官网, 信达证券研发中心

投资建议: AI 领域的发展速度超过传统制造业，2023 年来，Scaling Laws 开始发力，全球领域内开启了算力资源的“军备竞赛”。而在地缘政治等因素的激励下，算法层面的追赶终于由算力相对匮乏的中国团队唱响号角。我们认为，在通向 AGI 的道路上，算法创新和算力资源都是不可或缺的。早在 2024 年，我们就明确了推理市场要想实现繁荣必须降低推理成本的结论。如今，豆包、DeepSeek 研究团队的相关工作成功地大幅度降低了推理成本，推进了 AI 的产业化。DeepSeek 的成功，是开源模型相对闭源模型的一次飞跃。诚然，在十分有限的时间区间内，北美的科技大厂或把有限的精力放在算法层面，以充分挖掘自身算力资源的潜力。但是，这并不意味着对算力的投入将会停止。相反，AI 的产业化有望在中长期加速算力的投入，避免了以往的 AI 热潮最终因无法落地而走向失败的困局。

- **从推理市场看:** 推理有望快速放量，文生文、文生视频、图生视频的各类功能或快速迭代，并使得 AI 真正的理解物理世界，自动驾驶、人形机器人、AI 在各行各业掀起产业革命的奇点临近。
- **从训练市场看:** 一方面，训练模型的前沿探索仍需大量算力投入，且推理市场的放量，有望衍生新的模型探索方向。另一方面，如世界模型等方向的研究进展有望加速。但是，我们也看到大模型厂商的淘汰也将加速，无法超越开源模型的闭源大模型厂商或快速出清。

扎克伯格曾在 Meta 法说会上表示，“随着时间的推移，就像每个企业都有一个网站、一个社交形象和一个电子邮件地址一样，在未来，每个企业也将拥有一个客户可以与之交互的 AI 代理。我们的目标是让每个小企业，最终每个企业，都可以轻松地将其所有内容和目录提取到一个 AI 代理中，从而推动销售并节省资金。”这一时点正在逐步临近，从历史看，当计算机从实验室走向千家万户，并没有导致相关的企业没落，而是孕育了许多伟大的企业。目前看，AI 有望在经历这一过程，建议关注产业链相关标的。

表 1: 建议关注个股

板块	股票代码	股票简称	总市值 (亿元)	净利润 (亿元)			PE		
				2024E	2025E	2026E	2024E	2025E	2026E
AI 云侧	601138.SH	工业富联	4,261.69	246.17	308.54	356.03	17.31	13.81	11.97
	002463.SZ	沪电股份	797.95	25.45	33.66	40.43	31.35	23.70	19.74
	600183.SH	生益科技	724.44	18.40	23.83	28.92	39.37	30.40	25.05
	002916.SZ	深南电路	669.31	20.77	25.36	30.41	32.23	26.39	22.01
	300476.SZ	胜宏科技	459.21	11.76	18.45	22.98	39.04	24.89	19.98
	688256.SH	寒武纪	2,387.85	-4.59	0.27	4.78	-520.70	8,893.31	499.57
	688041.SH	海光信息	2,975.15	19.13	27.91	37.94	155.53	106.61	78.41
	300433.SZ	蓝思科技	1,296.05	39.62	55.17	68.42	32.71	23.49	18.94
	002600.SZ	领益智造	596.40	20.36	29.86	38.70	29.29	19.98	15.41
	002938.SZ	鹏鼎控股	938.79	36.14	45.23	51.37	25.98	20.76	18.27
AI 端侧	002384.SZ	东山精密	571.48	18.58	28.74	36.50	30.75	19.88	15.66
	688018.SH	乐鑫科技	303.21	3.46	4.58	5.98	87.51	66.23	50.74
	603893.SH	瑞芯微	689.09	5.02	7.38	10.19	137.22	93.33	67.63
	688608.SH	恒玄科技	476.59	3.87	5.81	7.96	123.05	81.98	59.90
	300458.SZ	全志科技	298.58	2.39	3.60	4.91	124.84	82.90	60.81
	603986.SH	兆易创新	846.01	11.19	16.56	21.22	75.63	51.10	39.86
	688099.SH	晶晨股份	351.40	8.14	10.97	14.04	43.16	32.02	25.02

资料来源: ifind, 信达证券研发中心 (ifind 一致预期, 截至 2025 年 1 月 27 日)

风险因素

宏观经济下行风险;

下游需求不及预期风险;

中美贸易摩擦加剧风险。

研究团队简介

莫文字，电子行业分析师，S1500522090001。毕业于美国佛罗里达大学，电子工程硕士，2012-2022 年就职于长江证券研究所，2022 年入职信达证券研发中心，任副所长、电子行业首席分析师。

郭一江，电子行业研究员。本科兰州大学，研究生就读于北京大学化学专业。2020 年 8 月入职华创证券电子组，后于 2022 年 11 月加入信达证券电子组，研究方向为光学、消费电子、汽车电子等。

杨宇轩：电子组分析师，华北电力大学本科，清华大学硕士，曾就职于东方证券、首创证券、赛迪智库，2025 年 1 月加入信达证券电子组，研究方向为半导体等。

王义夫，电子行业研究员。西南财经大学金融学士，复旦大学金融硕士，2023 年加入信达证券电子组，研究方向为存储芯片、模拟芯片等。

李星全，电子行业研究员。哈尔滨工业大学学士，北京大学硕士。2023 年加入信达证券电子组，研究方向为服务器、PCB、消费电子等。

分析师声明

负责本报告全部或部分内容的每一位分析师在此申明，本人具有证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师，以勤勉的职业态度，独立、客观地出具本报告；本报告所表述的所有观点准确反映了分析师本人的研究观点；本人薪酬的任何组成部分不曾与，不与，也将不会与本报告中的具体分析意见或观点直接或间接相关。

免责声明

信达证券股份有限公司（以下简称“信达证券”）具有中国证监会批复的证券投资咨询业务资格。本报告由信达证券制作并发布。

本报告是针对与信达证券签署服务协议的签约客户的专属研究产品，为该类客户进行投资决策时提供辅助和参考，双方对权利与义务均有严格约定。本报告仅提供给上述特定客户，并不面向公众发布。信达证券不会因接收人收到本报告而视其为本公司的当然客户。客户应当认识到有关本报告的电话、短信、邮件提示仅为研究观点的简要沟通，对本报告的参考使用须以本报告的完整版本为准。

本报告是基于信达证券认为可靠的已公开信息编制，但信达证券不保证所载信息的准确性和完整性。本报告所载的意见、评估及预测仅为本报告最初出具日的观点和判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会出现不同程度的波动，涉及证券或投资标的的历史表现不应作为日后表现的保证。在不同时期，或因使用不同假设和标准，采用不同观点和分析方法，致使信达证券发出与本报告所载意见、评估及预测不一致的研究报告，对此信达证券可不发出特别通知。

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，也没有考虑到客户特殊的投资目标、财务状况或需求。客户应考虑本报告中的任何意见或建议是否符合其特定状况，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测仅供参考，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人做出邀请。

在法律允许的情况下，信达证券或其关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能会为这些公司正在提供或争取提供投资银行业务服务。

本报告版权仅为信达证券所有。未经信达证券书面同意，任何机构和个人不得以任何形式翻版、复制、发布、转发或引用本报告的任何部分。若信达证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，信达证券对此等行为不承担任何责任。本报告同时不构成信达证券向发送本报告的机构之客户提供的投资建议。

如未经信达证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。信达证券将保留随时追究其法律责任的权利。

评级说明

投资建议的比较标准	股票投资评级	行业投资评级
本报告采用的基准指数：沪深300指数（以下简称基准）； 时间段：报告发布之日起6个月内。	买入 ：股价相对强于基准15%以上；	看好 ：行业指数超越基准；
	增持 ：股价相对强于基准5%~15%；	中性 ：行业指数与基准基本持平；
	持有 ：股价相对基准波动在±5%之间；	看淡 ：行业指数弱于基准。
	卖出 ：股价相对弱于基准5%以下。	

风险提示

证券市场是一个风险无时不在的市场。投资者在进行证券交易时存在赢利的可能，也存在亏损的风险。建议投资者应当充分深入地了解证券市场蕴含的各项风险并谨慎行事。

本报告中所述证券不一定能在所有的国家和地区向所有类型的投资者销售，投资者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、税收等方面咨询专业顾问的意见。在任何情况下，信达证券不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任，投资者需自行承担风险。