

电子

AI 行业点评：

DeepSeek 冲击全球，国产科技待重估

报告摘要

◆ 国产科技创新，DeepSeeK 树立 AI 性价比标杆

2025 年 1 月 20 日，幻方发布国产推理模型 DeepSeek-R1，在仅有极少标注数据的情况下，极大提升了模型推理能力。通过 DeepSeek-R1 的输出，蒸馏出 32B 和 70B 参数级别模型在数学、代码、自然语言推理等任务对标 OpenAI o1-mini 的效果。性能对标头部模型的同时，DeepSeek-R1 的服务定价为每百万 tokens 输入 1 元/输出 16 元，较 o1-mini 的价格（输入 11 元/输出 88 元）大幅降低。

◆ 开辟模型降本新路径，端侧 AI 普惠化可期

DeepSeeK 在基础模型训练和推理模型训练均有创新，有效克服了推高模型成本的 FP8 训练精度不足、高质量数据匮乏等困难，极大降低了训练和推理成本。DeepSeeK 的研究还表明，较大的基础模型发现的推理模式对于提高较小模型的推理能力至关重要。在大模型蒸馏技术的加持下，端侧小模型的普惠化指日可待。

◆ 全球科技巨头认可，重塑中国科技估值

全球科技公司都在探寻 DeepSeeK 可能带来的商机。美国科技巨头英伟达 NIM，微软 Azure，亚马逊 AWS 等智能云平台已上线部署支持用户访问 DeepSeek-R1 模型。我们认为，DeepSeeK 以创新的技术路线吸引了全球科技巨头，说明国产科技在开放创新的开源环境里具备强大的竞争力，有望涌现出更多的普适技术，我国本土科技企业的价值理应被重估。

◆ 建议关注

模型使用成本降低，AI 应用程序会越来越多，竞争将更加激烈，作为开发工具的 PaaS 服务的护城河更高。借助 DeepSeeK 引发的全球效应，我国本土科技公司的价值有望获得更高关注度。

投资评级

增持

维持评级

行业走势图



作者

刘牧野

分析师

SAC 执业证书：S0640522040001

邮箱：liumy@avicsec.com

建议积极布局 AI 工具，及端侧 AI 环节：

- 1) AI 服务：腾讯（港股）、阿里巴巴（港股）、金山云（港股）、第四范式（港股）；
- 2) AI 应用软件：泛微网络、金山办公、顶点软件；
- 3) 端侧 AI：兆易创新、中兴通讯、瑞芯微、恒玄科技。

◆ 风险提示：

AI 技术发展不及预期、AI 在应用端的表现不及预期、AI 受政策法规限制的风险、地缘政治的风险

正文目录

一、 国产科技创新，DeepSeek 树立 AI 性价比标杆	4
1.1 降本源自创新.....	5
1.2 蒸馏小模型表现出色，端侧 AI 普惠化可期	6
1.3 全球科技巨头认可，重塑中国科技估值	7
1.4 投资建议.....	8

图表目录

图 1 DeepSeek 与 OpenAI 模型性能对比	4
图 2 DeepSeek-R1 与 OpenAI o1 模型价格对比.....	4
图 3 DeepSeek 在安卓和苹果应用商店下载量位居前列	5
图 4 DeepSeek-R1 训练过程	6
图 5 DeepSeek-V3 性能对标国外模型.....	6
图 6 DeepSeek-V3 具高性价比	6
图 7 DeepSeek-R1 训练过程	7
图 8 RL 和蒸馏训练结果对比	7

一、 国产科技创新，DeepSeek 树立 AI 性价比标杆

2025 年 1 月 20 日，幻方发布国产推理模型 DeepSeek-R1，在仅有极少标注数据的情况下，极大提升了模型推理能力。通过 DeepSeek-R1 的输出，蒸馏出 32B 和 70B 参数级别模型在数学、代码、自然语言推理等任务对标 OpenAI o1-mini 的效果。

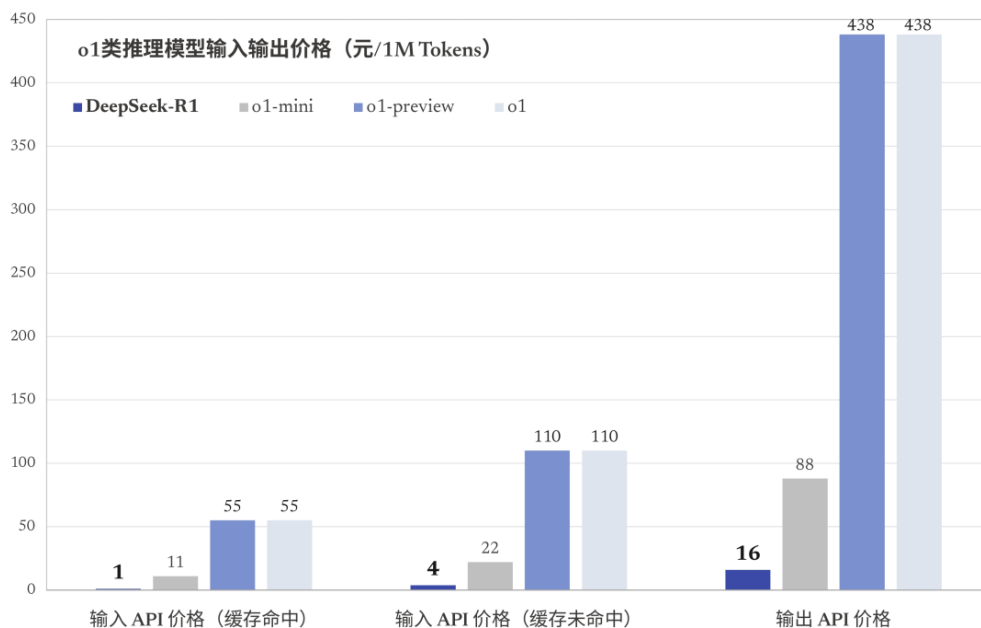
图1 DeepSeek 与 OpenAI 模型性能对比

	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：DeepSeek，中航证券研究所

性能对标头部模型的同时，DeepSeek-R1 的服务定价为每百万 tokens 输入 1 元/输出 16 元，较 o1-mini 的价格（输入 11 元/输出 88 元）大幅降低。

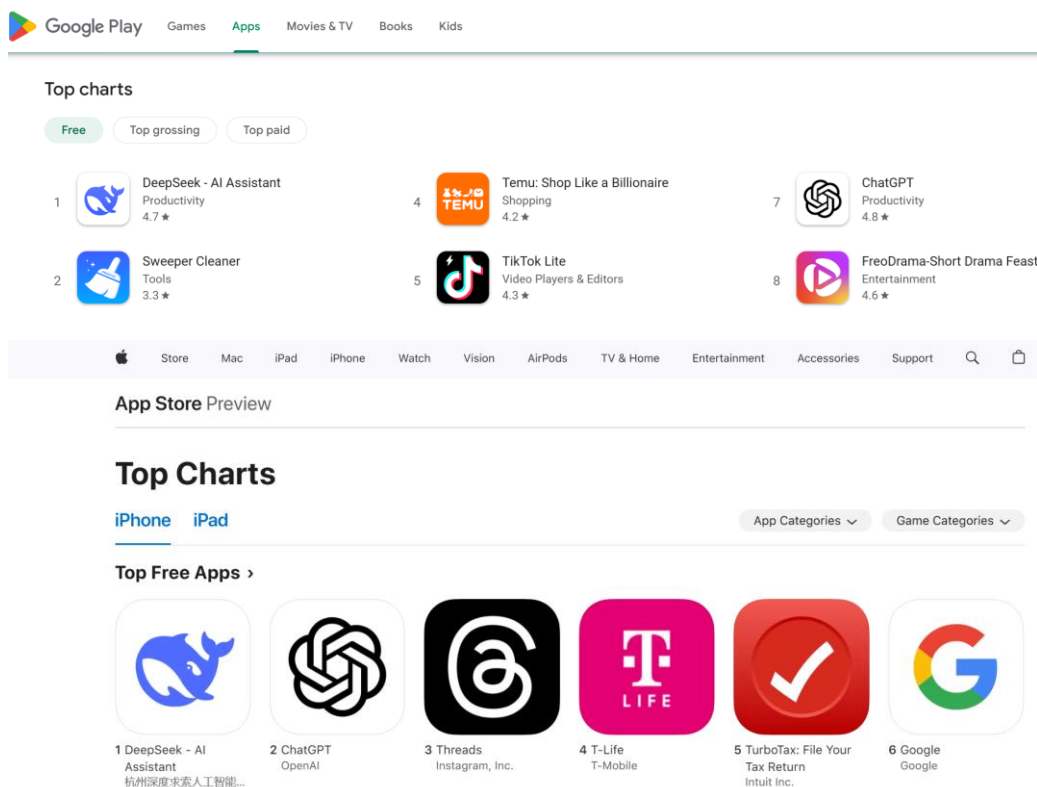
图2 DeepSeek-R1 与 OpenAI o1 模型价格对比



资料来源：DeepSeek，中航证券研究所

DeepSeek 是中国科技从追赶迈向创新的一大步，中国科技资产或将迎来一轮全面价值重估。作为国产大模型，DeepSeek 获得了国际范围的广泛认可，在安卓和苹果的应用商店获得了大量用户，用户下载量位居前列。DeepSeek 为业界开辟了大模型降本示范路径，将加速 AI 模型在硬件终端的渗透速度，并且获得了海外科技巨头的认可，说明中国科技创新能够在最前沿的科技浪潮中发挥全球影响力。

图3 DeepSeek 在安卓和苹果应用商店下载量位居前列

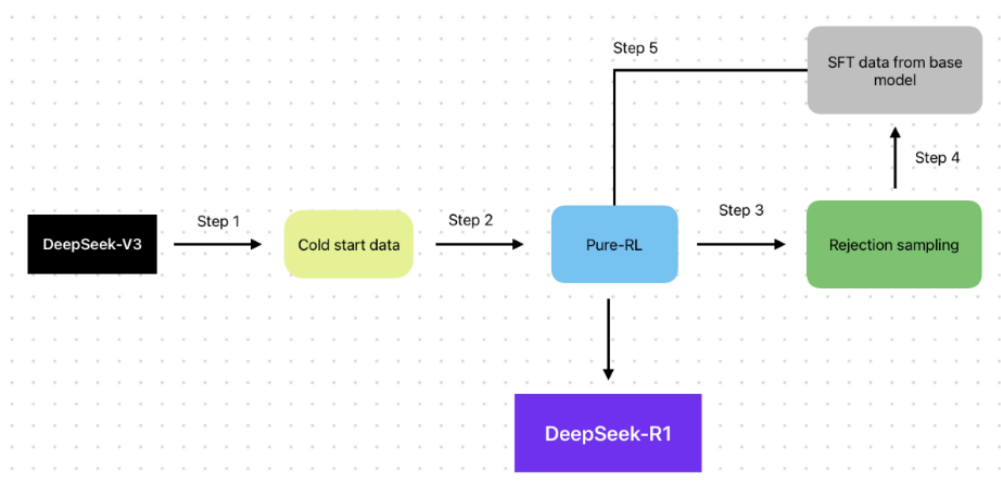


资料来源：google paly, app store, 中航证券研究所

1.1 降本源自创新

构建强大的 AI 推理能力一直面临着一个根本性的挑战：获取高质量的训练数据。DeepSeek-R1 开辟了“以智养智”方案来自动化获取数据，缓解了数据短缺的影响。R1 的训练架构为，在 DeepSeek-V3 基础模型之上，先利用少量高质量的“冷启动”（Cold Start）数据进行强化学习（RL）、拒绝抽样（Rejection Sampling）、数据微调（SFT），形成高质量数据集；接着用生成的高质量数据训练 DeepSeek-V3，重复以上步骤得到收敛的 R1 模型。这个方案的创新在于充分利用了推理任务的一个独特性质：答案的正确性往往可以通过程序自动验证。自动化的验证机制为强化学习提供了清晰的奖励信号，从而使模型能在较少人工标注的情况下不断改进，降低了模型对庞大高质量数据的要求。

图4 DeepSeek-R1 训练过程

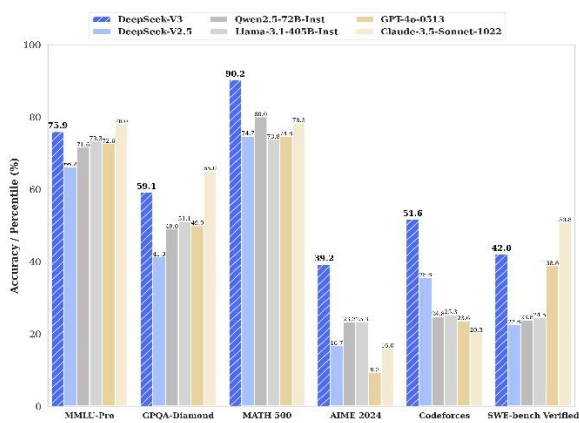


资料来源：vellum，中航证券研究所

R1 模型的简洁训练路径和低成本，还与 V3 基础模型的训练经验积累存在非常大的关系。DeepSeek-V3 是幻方于 2024 年 12 月 26 日发布的通用大模型，其大小和性能对齐海外领军闭源模型，但价格远低于同行。

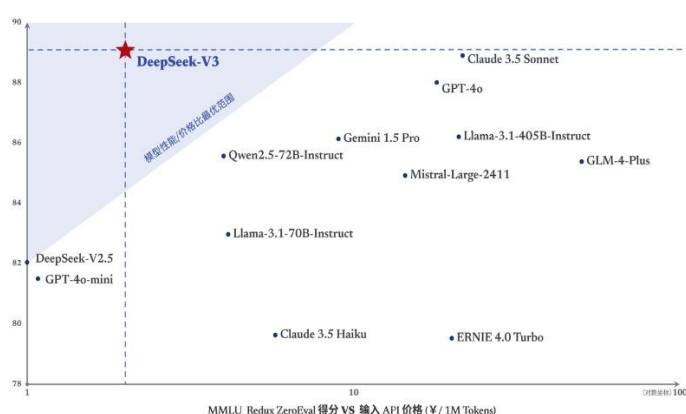
从历史上看，FP8 并未被广泛用于训练，因为梯度（在反向传播期间更新模型权重至关重要）在以如此低精度格式表示时通常无法收敛。DeepSeek-v3 通过引入一些细粒度量技术（例如逐块和逐块缩放）克服了这一挑战，在极大型模型上首次大规模实现 FP8 混合精度训练，并且持续优化模型架构，所以模型训练成本仅 557.6 万美元，总共使用了 278.8 万 H800 GPU Hours。相较之下，Meta 的 Llama 3 系列模型的计算预算则多达 3930 万 H100 GPU Hours。

图5 DeepSeek-V3 性能对标国外模型



资料来源：DeepSeekK，中航证券研究所

图6 DeepSeek-V3 具高性价比



资料来源：DeepSeekK，中航证券研究所

1.2 蒸馏小模型表现出色，端侧 AI 普惠化可期

DeepSeek 团队证明，通过完善的工程和算法设计，即使在有限的硬件和资源下

也能实现不错的 AI 性能。将较大模型的推理模式蒸馏为较小的模型，DeepSeek 获得了更好的性能。经过蒸馏的 14B 模型的表现远胜于先进的开源模型 QwQ-32B-Preview，而经过蒸馏的 32B 和 70B 模型在密集模型的推理基准上创下了新纪录。

图7 DeepSeek-R1 训练过程

Model	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	pass@1	cons@64	pass@1	pass@1	pass@1	rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717
OpenAI-o1-mini	63.6	80.0	90.0	60.0	53.8	1820
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9	1316
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1691
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1633

资料来源：DeepSeek，中航证券研究所

蒸馏像“临摹大师画作”，直接复现效果，适用于批量部署；RL 像“自己创作”，虽可能更好但费时费力，适用于做技术突破。使用 Qwen2.5-32B (Qwen, 2024b) 作为基础模型，DeepSeek-R1 的直接蒸馏 (DeepSeek-R1-Distill-Qwen-32B) 效果优于仅在其上应用 RL (DeepSeek-R1-Zero-Qwen-32B)。这表明，较大的基础模型发现的推理模式对于提高较小模型的推理能力至关重要。在大模型蒸馏技术的加持下，端侧小模型的普惠化指日可待。

图8 RL 和蒸馏训练结果对比

Model	AIME 2024		MATH-500	GPQA Diamond	LiveCodeBench
	pass@1	cons@64	pass@1	pass@1	pass@1
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9
DeepSeek-R1-Zero-Qwen-32B	47.0	60.0	91.6	55.0	40.2
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2

资料来源：DeepSeek，中航证券研究所

1.3 全球科技巨头认可，重塑中国科技估值

科技巨头纷纷接入，探寻 DeepSeekK 新机遇。DeepSeek 的技术路线有望整体降低 AI 技术的使用门槛，全球科技公司都在探寻其可能带来的商机。美国科技巨头英伟达 NIM，微软 Azure，亚马逊 AWS 等智能云平台已上线部署支持用户访问 DeepSeek-R1 模型。AMD 宣布已将新的 DeepSeek-V3 模型集成到 Instinct MI300X GPU 上。微软宣布推出针对 NPU 优化的 DeepSeek-R1 (已在 Azure AI Foundry 和

GitHub 上提供)，首批将面向搭载高通骁龙 X 芯片的 Copilot+ PC 推出，随后是英特尔酷睿 Ultra 200V 等平台。AI 编程工具 Cursor、AI 搜索引擎 Perplexity 等 AI 应用初创公司也上线了 DeepSeeK 服务。

我们认为，DeepSeeK 的低成本基础模型和端侧模型让全球认识到，对于云服务商和应用开发商而言，价格或成选择模型供应商的重要因素。DeepSeeK 以创新的技术路线吸引了全球科技巨头，说明国产科技在开放创新的开源环境里具备强大的竞争力，有望涌现出更多的普适技术，我国本土科技企业的价值理应被重估。

1.4 投资建议

模型使用成本降低，AI 应用程序会越来越多，竞争将更加激烈，作为开发工具的 PaaS 服务的护城河更高。借助 DeepSeeK 引发的全球效应，我国本土科技公司的价值有望获得更高关注度。

建议积极布局 AI 工具，及端侧 AI 环节：

- 1) AI 服务：腾讯（港股）、阿里巴巴（港股）、金山云（港股）、第四范式（港股）；
- 2) AI 应用软件：泛微网络、金山办公、顶点软件；
- 3) 端侧 AI：兆易创新、中兴通讯、瑞芯微、恒玄科技。

风险提示：AI 技术发展不及预期、AI 在应用端的表现不及预期、AI 受政策法规限制的风险、地缘政治的风险

公司的投资评级如下:

买入: 未来六个月的投资收益相对沪深 300 指数涨幅 10%以上。

增持: 未来六个月的投资收益相对沪深 300 指数涨幅 5%~10%之间。

持有: 未来六个月的投资收益相对沪深 300 指数涨幅-10%~+5%之间。

卖出: 未来六个月的投资收益相对沪深 300 指数跌幅 10%以上。

行业的投资评级如下:

增持: 未来六个月行业增长水平高于同期沪深 300 指数。

中性: 未来六个月行业增长水平与同期沪深 300 指数相若。

减持: 未来六个月行业增长水平低于同期沪深 300 指数。

研究团队介绍汇总:

首席: 赵晓琨 十六年消费电子及通讯行业工作经验, 曾在华为、阿里巴巴、摩托罗拉、富士康等多家国际级头部品牌终端企业, 负责过研发、工程、供应链采购等多岗位工作。曾任职华为终端半导体芯片采购总监, 阿里巴巴人工智能实验室供应链采购总监。长期专注于三大方向: 1、半导体及硬科技; 2、智慧汽车及机器人; 3、大势所趋的新能源。 分析师: 刘牧野 约翰霍普金斯大学机械系硕士, 2022 年 1 月加入中航证券。拥有高端制造、硬科技领域的投研经验, 从事科技、电子行业研究。

销售团队:

李裕淇, 18674857775, liyuq@avicsec.com, S0640119010012

李友琳, 18665808487, liyoul@avicsec.com, S0640521050001

曾佳辉, 13764019163, zengjh@avicsec.com, S0640119020011

分析师承诺:

负责本研究报告全部或部分内容的每一位证券分析师, 再次申明, 本报告清晰、准确地反映了分析师本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告中的具体推荐或观点直接或间接相关。

风险提示: 投资者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。

免责声明:

本报告由中航证券有限公司(已具备中国证券监督管理委员会批准的证券投资咨询业务资格)制作。本报告并非针对意图送发或为任何就送发、发布、可得到或使用本报告而使中航证券有限公司及其关联公司违反当地的法律或法规或可致使中航证券受制于法律或法规的任何地区、国家或其它管辖区域的公民或居民。除非另有显示, 否则此报告中的材料的版权属于中航证券。未经中航证券事先书面授权, 不得更改或以任何方式发送、复印本报告的材料、内容或其复印本给予任何其他人。

本报告所载的资料、工具及材料只提供给阁下作参考之用, 并非作为或被视为出售或购买或认购证券或其他金融票据的邀请或向他人作出邀请。中航证券未有采取行动以确保于本报告中所指的证券适合个别的投资者。本报告的内容并不构成对任何人的投资建议, 而中航证券不会因接受本报告而视他们为客户。

本报告所载资料的来源及观点的出处皆被中航证券认为可靠, 但中航证券并不能担保其准确性或完整性。中航证券不对因使用本报告的材料而引致的损失负任何责任, 除非该等损失因明确的法律或法规而引致。投资者不能仅依靠本报告以取代行使独立判断。在不同时期, 中航证券可发出其它与本报告所载资料不一致及有不同结论的报告。本报告及该等报告仅反映报告撰写日分析师个人的不同设想、见解及分析方法。为免生疑, 本报告所载的观点并不代表中航证券及关联公司的立场。

中航证券在法律许可的情况下可参与或投资本报告所提及的发行人的金融交易, 向该等发行人提供服务或向他们要求给予生意, 及或持有其证券或进行证券交易。中航证券于法律容许下可于发送材料前使用此报告中所载资料或意见或他们所依据的研究或分析。

联系地址: 北京市朝阳区望京街道望京东园四区 2 号楼中航产融大厦中航证券有限公司

公司网址: www.avicsec.com

联系电话: 010-59219558

传 真: 010-59562637