



人工智能文集

第二十四集

中国开源软件推进联盟

China Open Source Software Promotion Union

目 录

- 一、中国 DeepSeek 模型正在改变全球 AI 格局.....Geoffrey Hinton
(AI 大师辛顿应普林斯顿读书会专访发表讲话, 陈伟节录)
- 二、DeepSeek 引发的 AI 发展路径思考.....李国杰
- 三、再谈 DeepSeek-v3, -R1 降低训练成本问题.....陆首群
- 四、DeepSeek: “国运级创新”, 凭啥?谭中意
- 五、DeepSeek 与 Grok3 的对比.....陈钟摘录
(从《关于 Grok3 和 DS 背后苦涩教训引发思考》一文中摘录)
- 六、点评 DeepSeek-R1 与 Grok3 生成式 AI 排行榜.....COPU
- 七、从 DeepSeek R1 看大模型技术进展和趋势.....倪贤豪
- 八、点评萨姆-奥特曼.....COPU
- 九、阿里云千问 QwQ-32B 推理模型性能媲美 DeepSeek R1!宋逸群
- 十、再谈谈我对 Manus 的见解.....陆首群
- 十一、DeepSeek 提前发布 R2.....谭中意摘录

中国 DeepSeek 模型正在改变全球 AI 格局

Geoffrey Hinton

(AI 大师辛顿应普林斯顿读书会专访发表讲话，陈伟节录)



AI 大师辛顿（Geoffrey Hinton）在今年首次接受由普林斯顿读书会主持的一次深度专访，针对中国 AI 代表作 DeepSeek 给出了令人振奋的评价，引发业内广泛关注。

主持人：你也看到像中国那边的 DeepSeek，研发成本大致在 500 万至 600 万美元之间，可以做出接近高的成果。

辛顿：这一事实正在改变全球 AI 格局。在全球 AI 技术竞赛日益白热化的当下，DeepSeek 等中国模型正以惊人的性价比挑战西方 AI 巨头的技术优势，展示了中国在 AI 领域的创新能力。

在谈到近来 Transformer 和大规模语言模型开始席卷整个 AI 领域，这时西方普遍感到算力不足，需要加大投资配置巨大的算力资源等。

主持人：我知道英国、美国、加拿大也有一些安全研究所正朝这个方向努力，可如果你去看它们能使用的算力资源，其实很有限。英国那家研究所也许能微调一个大模型，但要想从头训练大模型就力不从心了，只有大公司才有这个本事。

辛顿：是啊，不过你也看到像中国那边的 DeepSeek，研发成本大致在 500 万至 600 万美元之间，可以做出接近高的成果，而且他们确实把它开源了。可是关于“开源”，我觉得也有值得商榷的地方。有些人说是“开放源代码”，但其实并不一样。传统的“开源”是开放代码，比如 UNIX 内核那样，人们可以查看代码发现 bug，而把权重放出来，人们并不会去看一行行权重，说“哦，这个权重好像太大了”。而当你把权重都公开，坏人只要一点点资源就能对模型进行“微调”，让它干各种坏事。所以我认为，不应该公开大模型的权重。这就好比把“核裂变材料”公诸于众。如果你问，“为什么我们每个人都没有原子弹呢？为什么每个小镇没有自己的核弹呢？”答案是要提炼铀的同位素非常难，可如果你把提炼好的核材料发给大家，那做炸弹就很简单了。对大型语言模型来说，“炼铀”的阶段就是预训练那个基础模型，而那需要海量资源。也许比以前要少了，但依然不少。可一旦你把训练好的权重公布出来，坏人只需要几百万美元就能让它干各种坏事。把权重交出去简直太疯狂了。

DeepSeek 引发的 AI 发展路径思考

李国杰

1 为什么 DeepSeek 会引起全球性的科技震撼

DeepSeek 的横空出世是人工智能 (AI) 发展史上新的标志性事件。7 天之内 DeepSeek 的用户增长超过 1 亿，创造了用户增长速度新的世界纪录。与此同时，芯片巨头公司英伟达 (NVIDIA) 的股价单日暴跌 17%，市值缩水 5890 亿美元，创下美国上市公司单日最大损失纪录。DeepSeek 的崛起，打破了“高算力和高投入是发展人工智能唯一途径”和“集成电路制程优势=人工智能技术霸权”的迷信，引领人工智能行业进入以算法和模型架构优化为主，同时高度重视数据质量与规模、理性提高算力的新时期。同时，DeepSeek 的崛起也标志着中国科技公司从“追赶者”变为“规则改写者”，在全球最关注的人工智能领域，以颠覆性的创新开始挑战西方在 AI 领域的霸权。

全球人工智能龙头企业纷纷拥抱 DeepSeek，凸显其不可抗拒的影响力。微软最早宣布将 DeepSeek R1 模型添加到其云平台 Azure AI Foundry，开发者可用于构建基于云的应用程序和服务。亚马逊云科技 (AWS)、英伟达、超威半导体 (AMD) 等公司相继宣布在其 AI 服务平台上部署 DeepSeek V3 和 R1 模型。不管某些国家政府如何抵制，少数媒体如何恶意中伤，上亿用户和众多大公司根据 DeepSeek 的性价比和亲身体验做出选择，主动融入 DeepSeek 生态。DeepSeek 推出的高效率、低成本的推理模型和开源商业模式，将引领人工智能行业新潮流。

DeepSeek 的 V3 和 R1 模型广受欢迎，首先是因为其在模型算法和系统软件层次都有重大创新。DeepSeek-V3 的模型参数量高达 6710 亿，但由于采用了自主研发的混合专家模型 (MoE) 架构，每一层有 256 个细分领域的路由专家和 1 个共享专家，每次调用只激活约 370 亿个参数，显著降低了训练计算成本。DeepSeek 改进的多头潜在注意力机制 (MLA)，减少了键值缓存开销，把显存占用降到了其他大模型的 5%~13%，极大提升了模型运行效率。DeepSeek-R1 模型摒弃了传统的监督微调 (SFT)，开创性地提出群组相对策略优化 (GRPO)，直接通过强化学习从基础模型中激发推理能力，大幅降低了数据标注成本，简化了训练流程。DeepSeek 揭示了一个真相，即推理模型的开发比想象中更为简单，各行各业都可以做。DeepSeek 的这些发明并非首次提出来的原始创新，但 DeepSeek 通过艰苦的努力把技术做到极致，在前人公开成果基础上，登上新的技术高峰。

第 3 波人工智能兴起后，美国政府、AI 龙头企业和投资界形成一个基本信念：发展人工智能需要高算力，而目前实现人工智能计算性能最高的芯片就是英伟达的 GPU。因此，美国认为只要控制 GPU 的销售，就能在人工智能领域独霸全球。特朗普在就职典礼第 2 天就签署法案，启动星际之门计划，投资 5000 亿美元，打造人工智能的基础设施。

可见，美国政府是把夯实人工智能的算力基础当成维持其全球领导地位的关键。以英伟达为代表的美国 AI 龙头企业，有真老虎的一面，也有纸老虎的另一面。初生牛犊不怕虎的年轻的中国科技工作者，在纸老虎上戳了一个洞，让世人看清了原来这只老虎也没有那么可怕。DeepSeek 震撼全球就是揭示真相的威力。

2 “规模法则（Scaling Law）”是否已遇天花板

2020 年 1 月，OpenAI 发表论文《神经语言模型的规模法则》（Scaling Laws for Neural Language Models），提出规模法则：“通过增加模型规模、数据量和计算资源，可以显著提升模型性能。”在 AI 领域，规模法则被一些人认为是“公理”，俗称“大力出奇迹”，OpenAI 等龙头企业和美国的 AI 投资界把它当成制胜法宝。但规模法则不是像牛顿定律一样经过无数次验证的科学定律，而是 OpenAI 等公司近几年研制大模型的经验归纳。从科学研究的角度看，属于一种对技术发展趋势的猜想；从投资的角度看，属于对某种技术路线的押注。人工智能是对未来技术的探索。技术路线存在多种可能，人工智能本身也存在多元化的目标，探索的道路上有很多高山需要去攀登，攀登一座高山的路径也不止一条。把一种信仰或猜想当成科学公理，不是科学的态度。近几年大模型训练的实际效果表明，要获得大模型性能的线性增长，必须在模型规模、数据量和算力投入上高指数性的增长，几个月就翻一倍。从 GPT-3 到 GPT-4，参数规模增加约 10 倍，用于训练的 GPU 数量增加了近 24 倍，总计算量增加了近 70 倍。任何投入的高指数性增加都不可能是长久的，民航大飞机的速度提高、集成电路的主频提高等都是到适当的时候就停止了，大模型也应该不会破例。

鼓吹“Scaling Law”的人，常以强化学习之父理查德·萨顿（Richard S.Sutton）的文章“苦涩的教训”作为追求高算力的依据：“研究人员曾一次又一次试图通过精巧的工程设计来提升性能，但最终都败给了简单粗暴的‘加大算力’方案，历史证明，通用方法总是在 AI 领域胜出。”但是，萨顿本人这两年对“Scaling Law”做了深刻反思。他指出，虽然 Scaling Law 在提升模型性能方面确实有效，但它并不是解决所有问题的万能钥匙。AI 系统不仅需要具备强大的计算能力，还需要具备持续学习、适应环境、理解复杂情境等能力，这些能力往往难以通过简单地增加算力来实现。

但现在就说规模法则已经走到尽头，也没有根据。与人脑的神经连接复杂性相比，现在的人工神经网络至少还有上百倍的差距。继续扩大神经网络的规模和增加训练的数据量，是否还能取得与投入相称的回报，要看今后的实际效果。但 GPT-5 迟迟不能问世，可能说明规模扩张的效果已经不太明显。图灵奖得主杨立昆（Yann LeCun）和 OpenAI 前首席科学家伊利亚·苏茨克维（Ilya Sutskever）等直言，规模法则已触及天花板。

DeepSeek 的出现，逼迫 AI 界严肃地思考这一技术发展路线问题：是继续投入巨资追求高算力，还是另辟蹊径，在算法优化上下更多功夫？DeepSeek 的问世标志着人工智能训练模式从“大力出奇迹”的外延式发展阶段，转向集约化系统优化的内涵式发展阶段。

DeepSeek 的成功并没有否定算力在人工智能发展中的重要作用。实际上，由于用于推理的设备比训练设备多得多，推理所需要的算力将来会成为主要需求。但绿色发展是必须遵循的大原则，降低人工智能所需要的能耗一定是科技界的重要目标。

3 发展“通用人工智能”（AGI）应选择什么道路

“通用人工智能”是一个模糊的没有形成广泛共识的术语。OpenAI 公司追求的通用人工智能（artificial general intelligence, AGI）是其中的一种，指的是 AI 在多个领域以人类水平处理复杂问题的能力。人工智能界有一个莫拉维克悖论：“复杂的问题是易解的，简单的问题反而是难解的。”从这个角度看，能解复杂问题的人工智能不一定是通用人工智能。许多人认为，能够应对设计者预料之外的情况，才叫“通用”。因此，人工智能学术界更关注智能系统持续学习，自我改进的能力。人工智能的通用性不仅表现在对语言的处理上，还包括像人一样基于常识和日常经验与外部客观世界互动的能力。

人工智能是对人类智能某一个方面的再现和超越。在科学技术领域，所谓“通用”一定是相对的，有一定的条件或范围。我们要认识人工智能的局限性，不能盲目追求能够解决所有问题的人工智能。重点还是要根据实际需求，将相对通用的人工智能技术落地到各个行业，让一定范围内的人工智能技术见到实效。实现通用智能是一个渐进过程，不会因某项技术的发明就突然到来。人工智能的通用性已经比前两波有明显的提高，但在某些应用中通过图灵测试只是阶段性成果，目前的技术离真正的通用智能还有较大差距。

究竟如何才能实现通用人工智能，现在还没有结论。DeepSeek 和 OpenAI 都以发展“通用人工智能”为目标，但走的路径不一样。OpenAI 公司相信 Scaling Law 正在尽量扩大模型规模，希望先做出通用的基础模型，再“蒸馏”出各行业可使用的行业垂直模型，走的是“由通到专”的路。除降低通用大模型的训练成本外，如何在保持泛化能力的同时，提高在特定领域或任务上的性能和效率，仍是需要解决的问题。与之相反，DeepSeek 走的是“由专到通”的人工智能发展之路，试图在模型算法和工程优化方面进行系统级创新，为在受限资源下探索通用人工智能开辟新的道路。所谓“混合专家模型”就是集小智为大智，集专智为通智。“小而精”的模型将人工智能的重点发展方向从面向企业的 to B 引向更贴近消费者的 to C，从广泛覆盖的“横向通吃”引向深入钻研的“纵向做精”，让更多的中小企业参与，可能会创造更大市场空间。但整合多个专用模型为通用模型也需要解决诸多技术和工程问题，如模型间的接口、数据格式的统一、训练时的负载平衡等。

通用与专用的竞争是技术发展的普遍现象，集成电路领域就有“通用”与“专用”10 年交替演化的“牧本周期”。“由通到专”和“由专到通”的人工智能，究竟哪条路能走通，还要等历史下结论。也许最后的结果是通专融合，专用多模型的“垂直做精”与通用大模型的“横向扩展”形成互补，共同构建起智能时代的产业新生态。

4 发展人工智能应该追求高算力还是高算效（高能效）

科技界公认图灵是人工智能的奠基人，因为他提出了可以用计算来模拟人类智能的科学假说。他的论文暗示，计算等价于智能。迄今为止，人工智能取得的成果几乎都离不开计算，大模型的出现将算力的作用抬高到前所未有的高度。我们需要认真地想一想，高算力是不是人工智能的本质要求？

发展人工智能的初始动机是模拟人脑，自然界进化了数百万年的人脑是一个计算效率和能效极高的计算装置，功耗只有 20W 左右。人脑的极低功耗是因为采取了分布式的模拟计算。目前计算机的高能耗是因为采用软硬件分离的数字计算。深度学习的奠基人辛顿（Hinton）教授最近提出“凡人计算”的新研究方向，采用与人脑一样的存算一体模拟计算方式，颠覆了硬件与软件分离的传统计算模式。这类研究追求的是计算的高算效和高能效，从长远来讲，是发展人工智能的正确方向。

DeepSeek 发布以后，斯坦福大学华裔科学家李飞飞指导的团队，以阿里通义千问（Qwen）模型为基础，通过“蒸馏”谷歌的 AI 推理模型 Gemini 2.0 Flash Thinking Experimental，并结合 SFT 技术，使用 16 个英伟达 H100 GPU 进行了 26min 的训练，以不到 50 美元的云计算费用成功训练出 s1 模型，其性能超过了 Open AI 的 o1-preview 模型。这种低训练成本的模型的通用性可能不如大公司的模型，但令人吃惊的低成本可在某些应用上做到与高端模型相媲美的性能，显示出人工智能的低成本化还有巨大的提升空间。

低成本是技术普及的基本要求，蒸汽机、电力和计算机的普及都是其成本降低到大众可以接受时才做到的，人工智能肯定也会走这条路。目前，盲目地追求高算力导致人工智能的成本居高不下，阻碍了人工智能技术的大规模普及。DeepSeek 不仅是技术突破者，更是规则重构者，开辟了一条低成本发展人工智能的可行之路。DeepSeek 的崛起说明 AI 不再局限于简单堆砌算力，而是进入了以追求高算效和高能效为主的新阶段。

5 “开源”为什么有这么大的威力

过去几年，开源大模型的性能始终与龙头企业的闭源大模型有一代以上的差距，这一次 DeepSeek 的性能追上了闭源模型，大大增强了开源社区的信心。图灵奖得主杨立昆认为，“对 DeepSeek 崛起的正确解读，应是开源模型正在超越闭源模型”。这一评价十分中肯，因为改变 AI 发展模式比单项技术的突破更重要。

虽然 OpenAI 为代表的生成式人工智能红红火火，但多数企业不敢将本公司的数据交给私有 AI 平台生成自己的垂直模型，因为害怕提交的数据泄露自己的技术机密。人工智能之所以在各个行业难以落地，这可能是一个根本性的原因。DeepSeek 的彻底开源模式破解了这一难题，现在全世界的企业和用户都可以将 DeepSeek 提供的小而精的模型下载到本地，即使断网也可以“蒸馏”出高效率的垂直模型，真正实现了技术的民主化。

长期以来，美国的 AI 龙头公司夸大开源 AI 的安全风险，力图通过监管抑制开源 AI。事实上，开源模型对于全球 AI 供应链至关重要，发展中国家尤其需要开源 AI 技术。若美国继续在这一领域设置障碍，中国就有望在开源 AI 全球供应链中占据核心地位，从而使更多企业转向中国企业而非美国企业的技术方案。

真正的 AI 竞争，不仅仅是技术和模型的竞争，更是生态系统、商业模式，以及价值观的竞争。开源模型让每个开发者都能轻松调用强大 AI 工具，不再受大公司的约束，AI 的进化速度将会明显提升。DeepSeek 的开源战略将向历史证明：在这场 AI 竞赛中，谁拥抱开源，谁就能赢得未来。

6 中国是否已具有在人工智能上引领全球的实力

有人说 ChatGPT 是 0 到 1 的突破，而 DeepSeek 只是 1 到 N 的扩展，这种看法不符合人工智能发展的历史轨迹。人工智能是一个没有严格定义的研究领域，没有智能和不智能的 0 到 1 界限，只有智能化水平不断提高的发展过程。长期以来，中国人工智能领域的高技术企业大多重视应用创新和商业模式创新，追求的目标是快速盈利，很少参与核心技术创新。随着经济的发展和技术的积累，中国的企业已经开始具备原创的能力。DeepSeek 的一鸣惊人可能是一个分水岭，标志着中国 AI 产业开始从“技术跟跑”向“技术并跑和领跑”迈进。

应当承认，中国在人工智能的基础研究和核心技术上与美国仍然有差距。尽管在人工智能领域，中国的论文发表总数和专利授权数量超过美国，但是引用最多的源头性论文大多出自美国，美国也是顶尖 AI 模型的主要来源国。斯坦福大学发布的《2024 年人工智能指数报告》显示，2023 年，美国有 61 个较著名的 AI 模型，中国只有 15 个。近几年中国在 AI 领域快速追赶，进步速度喜人。根据《日本经济新闻》对 2020—2024 年神经信息处理系统大会（NIPS）等 3 个机器学习顶级会议的统计，在 3 万多篇发表的论文中，中国作者有 8491 人（美国 14766 人），过去 4 年中国作者增长了 8 倍。

人工智能不同于资本密集型和经验积累型的集成电路产业，不仅要“烧钱”，更要“烧脑”，本质上是拼人的智力的新兴产业。因此人工智能产业具有明显的不对称性，一个具有 100 多个聪明头脑的小企业就可以挑战市值上万亿的龙头企业。DeepSeek 只是中国具有潜力的人工智能企业之一。近日，《麻省理工科技评论》刊发了一篇题为《关注 DeepSeek 之外的 4 家中国人工智能初创公司》的报道，指出阶跃星辰（Stepfun）、面壁智能（Model Best）、智谱 AI（Zhipu）、无问芯穹（Infinigence AI）4 家企业同样展现出不逊于 DeepSeek 的技术实力与全球竞争力。DeepSeek 脱颖而出以后，人们盛传“杭州 6 小龙”的故事（深度求索、宇树科技、游戏科学、云深处、群核科技、强脑科技 6 家初创 AI 公司）。迄今为止，中国共有 52 家人工智能领域的独角兽企业，约占全球人工智能独角兽企业的 18%。这说明在 AI 领域，中国已经有一批创新型小企业进入世界前列，开始展现引领全球的实力。

DeepSeek 的成功说明，在人工智能的发展中，算法优化和系统级工程优化不可或缺，优秀工程师发挥着至关重要的作用。良好的工程教育体系，庞大的工程师队伍是中国的一大优势，我们要充分发挥这一优势。进入并跑状态以后，就不要过于在意与美国相差几个月，可以你做你的、我做我的，比谁能发现正确的研究方向。年轻一代正在成为科研的主力，我们要有信心在人工智能的研究和应用上走在美国前面。

7 中国实现人工智能自立自强要如何发力

实现人工智能自立自强，不仅要靠国家的顶层规划和充分的资金支持，更要做好人才的使用培养和产业生态的构建，克服重重困难的前提是要有自信心。DeepSeek 成功的前提是其创始人梁文锋的自信。他在接受采访时说：“中国的 AI 不能永远做跟随者，必须有人走在前沿。OpenAI 并非神一般的存在，他们不可能永远领先。”80 后、90 后的中国年轻人已经开始平视美西方国家，具有“敢为天下先”的勇气和自信，他们是中国科技自立自强的希望。DeepSeek 的人才聘用模式打破了传统，梁文锋选择了一条与众不同的用人策略。他拒绝经验丰富的人才，而选择初出茅庐的年轻人。招聘时工作经验超过 8 年者，直接拒聘；超过 5 年，需要特别出色才能入选。DeepSeek 团队成员几乎全是国内顶尖高校的应届毕业生或博士实习生。真正的创新往往来自那些没有包袱的人，DeepSeek 重用极致热情、富有好奇心的年轻人，而不是习惯用经验找答案的人，这种用人理念为公司带来了令人惊奇的创新动力，也对传统的中国教育模式和人才聘用模式提出了警示。

要实现人工智能自立自强，最困难的是构建自主可控的产业生态。英伟达公司的“护城河”不是 GPU 芯片本身，而是统一计算设备架构（compute unified device architecture, CUDA）软件生态。DeepSeek 冲击了 CUDA 生态，但没有完全绕过 CUDA，其生态壁垒仍然存在。从长远来讲，需要开发一套比 CUDA 更优秀的自主可控的 AI 软件工具系统，重构 AI 软件生态。实现这个目标需要周密的规划和长期努力，相关部门应该下决心组织全国的开发力量，充分调动上下游企业的积极性，完成这件大事。

资金投入不是决定 AI 成败的唯一因素，但最近几年中国投资市场规模急剧萎缩值得警醒。CB Insights 数据显示，2023 年，美国的 AI 投资达到 672 亿美元，是中国 AI 投资的 8.7 倍。这一年美国的 AI 投资实现了 22.1% 的增长，而中国 AI 私人投资下降了 44.2%。其中，在生成式 AI 私人投资上，2023 年美国总投资额达到 224.6 亿美元，中国仅为 6.5 亿美元。风险投资和私募基金对于支撑科创产业极为重要，他们为创新起到了资金池和担保的作用。中美科创市场曾经并驾齐驱，但到 2023 年，中国科创投资额仅相当于美国的 8%。尽管美国投资界追捧“大算力”，存在一定的泡沫，但正常的金融支持是发展 AI 的必要条件。政府和资本界要合力构建一个健康的科创金融生态，为创新提供必备的动力，这样才会有更多 DeepSeek 出现。

产业生态的形成也要靠市场牵引。国家要通过政策引导，鼓励在 PC、手机和物端设备上推广 AI 应用，增强国产 GPU、CPU 和国产软件的市场占有率。要高度重视芯片设

计和大模型的开源战略，争取中国在全球人工智能的开源系统中起到主导作用。我们需要在有限算力条件下，通过算法和软件的协同创新，充分发挥硬件的极致性能，挖掘所有可能的优化空间。国内的人工智能模型与美国的水平十分接近，我们要做好算力资源与人工智能平台的优化适配，力争中国的人工智能科研和应用走在世界前列。

再谈 DeepSeek-v3, -R1 降低训练成本问题

陆首群

深度求索（DeepSeek）公司梁文锋团队向世人发表重大的人工智能（AI）研发成果：DeepSeek-v3,R1，震撼硅谷，继而震撼世界！他们发表 DeepSeek 的重大意义有三：1）成功地解决发展 AI 的路径问题：他们否定了美西方以巨额投资、高成本、堆叠资源，闭源、低效率，只供少数寡头发展 AI 的旧路，创造性地走出一条以低投资、低成本、简约资源，开源、高效率，发展 AI 的新路；2）打击美西方称霸 AI 领域的野心；3）激励中华奋起发展 AI 推动新科技革命的雄心壮志。

在刘明、章文嵩、刘锋等一众朋友的协助下，我终于有机会了解到 DeepSeek-v3,R1 大模型训练成本下降的情况，现在再次对其作出说明（并与其对标的 GPT-4，GPT-4o 做对比）。

造成 Deep Seek-v3,R1 训练成本下降的因素有：1）“多头潜在注意力架构（MLA）”的影响，2）“混合专家架构（MoE）”的影响，3）强化学习的目标函数（GRPO）的影响，4）发挥数学推导、代码生成、物理建模和工程突破能力（STEM）的优势。

DeepSeek-v3 大模型训练成本为 560 万美元（训练消耗 280 万 GPU 小时，以每个 GPU 小时 2 美元计费，则其训练成本为 560 万美元），与其对标的 GPT-4 大模型的训练成本为 7800 万美元，DeepSeek-v3 训练成本下降至 1/14(7%)；DeepSeek-R1 大模型的训练成本为 600 万美元，与其对标的 GPT-4o 大模型的训练成本约为 1.2 亿美元，DeepSeek-R1 训练成本下降至 1/20(5%)。

(DeepSeek-R1 的定价权约为 GPT1o 的 1/30，供参考)

归纳起来，DeepSeek-v3,R1 大模型的训练成本较其对标的 GPT-4，GPT-4o 大模型下降至 1/14~1/20(7%~5%)。

(DeepSeek-v2 大模型训练成本比 GPT-4Turbo 下降至 1/70)，1.43/6,供参考)

从 DeepSeek-v2 开始，由高效推理的核心“多头注意力架构 MLA”升级原来“标准的多头注意力

架构 MHA", 把显存占用降低 5%~13%, 取 10% (相当于训练成本下降 10%); 同时其独创的 DeepSeek MoE Sparse 架构可将计算量降到极致, 或把训练成本砍了 40%。如此说来, DeepSeek-v2 较之 DeepSeek-v1 可将训练成本下降 10%+40%=50%。对于 DeepSeek-v3, R1 的“混合专家模型 MoE”, 其计算量降低相当于降低训练成本, 在理论上, 计算量降至全密集的激活参数量/总参数量: 37B/671B=5.5%。

(在 DeepSeek-v3 的技术报告中谈道: 对于 DeepSeekv3-MoE 而言, 一次激活 37B/671B 的参数比例为 1/18.16 (5.5%), 该技术报告说明: 这是推理和训练上最省成本的部分)

如按首先激活细分的“专家系统”来考察 DeepSeek-MoE 降低训练成本问题: DeepSeek 16B 参数模型采用“混合专家”架构, DeepSeekv3-MoE 16 B (参数) 模型, 在保持高性能同时可大幅降低计算成本。MoE 架构包含 256 个专家, 每次根据输入选择激活其中 8 个专家 (即每个 token 分配 8 个专家), 一次激活 $8/256=1/32$, 即 DeepSeekv3-MoE 可将训练成本降至 1/32(3%)。

归纳起来, DeepSeek-MoE 在首先激活“参数量”时, 可使训练成本降至 1/18(5.5%), 如首先激活“专家”时, 可使训练成本降至 1/32 (3%)。

对于群组相对策略优化 GRPO (直接通过强化学习, 从基础模型中激发推理能力), 是一种如何减少训练资源要求的关键算法, 可大幅降低了数据标注成本, 简化训练流程并降低训练成本, 其降低幅度可从实践中获得 (GRPO 在训练过程中的收获往往用于补偿 MoE 的损失部分)。

发挥 STEM 的优势 (在数学推导、代码生成、物理建模、工程突破、提高决策能力等方面比较突出), 从操作实践中取得降低训练成本的数据。因具有 STEM 优势导致 DeepSeek 大幅降低训练成本的现象, 在当时测试中引起了硅谷 AI 专家的关注。

下面谈谈清华方案: 清华大学 Cache、AI 团队与趋境科技合作开发 KTransformers, 一块单张 RTX4090 显卡就能流畅运行 7B 模型, 花费仅 2 万美元。

我们的点评: 清华方案做不了训练, 但确实能把推理成本降到 1~2 万美元这个数量级。

对于 DeepSeek 训练成本下降情况综合分析:

对于 DeepSeek-v3,

企业推荐可降至：1/14（7%），

+MoE 可降至：1/18（5.5%），

+MoE（如首先激活专家系统）可降至 1/32（3%）。

对于 Deepseek-R1,

企业推荐可降至 1/20（5%），

+MoE（如首先激活专家系统）可降至：1/32（3%）。

DeepSeek：“国运级创新”，凭啥？

谭中意

前言：

春节前后，要说国内外科技圈最火热的名字，DeepSeek 绝对算头一个。“开源大模型之光”、“AI 领域新星”、“有望比肩 OpenAI”…… 各种赞誉纷至沓来，甚至有人直接将其冠于“国运级创新”的说法。“国运级创新”？ 这个帽子可实在不小。但是有人会质疑，凭啥？本文则试着从产业共识的角度来分析，笔者认为它确实担得起。

1. DeepSeek，是国运级创新！

要探寻“DeepSeek 是国运级创新”这一说法的源头，最早的说法来自游戏圈内极具影响力的人物——现象级 3A 游戏《黑神话：悟空》的制作公司游戏科学创始人冯骥。他在新浪微博上就曾公开表示：“DeepSeek 可能是个国运级别的科技成果”。冯骥以其独到的眼光和对技术趋势的敏锐洞察力而著称，他的这番评价，无疑引发了广泛关注。下图是他的微博截图。



Yocar-冯骥

游戏科学创始人、CEO，《黑神话：悟空》制作人

关注 442 | 粉丝 59万

2025/1/26 23:32

年前还是认真推下DeepSeek吧。

V3已经用了一个月，但R1仅仅用了5天，不得不过来发这条微博。

先讲一句暴论：DeepSeek，可能是个国运级别的科技成果。

公众号 · 开源老DJ

国内安全领域的领军人物，360 集团创始人周鸿祎也表达了对 DeepSeek 的有力支持。在今年两会期间接受《新京报》记者关于“人工智能技术开源”话题的采访时，他“举双手赞同”冯骥关于 DeepSeek “国运级别的科技成果”的评价。他认为开源模式形成了巨大的虹吸效应，一旦形成气候，将彻底战胜闭源。凭借更低的成本和更开放的技术普惠路径，中国在 AI 应用规模和渗透率上比美国更高。周鸿祎作为在 IT 行业深耕多年的资深人士，他的认可更显分量。



（图片来自新京报）

著名技术社区 CSDN 的高级副总裁李建忠今年也撰文《DeepSeek 关键技术创新及对 AI 生态的影响》，从技术生态的角度深入分析了 DeepSeek。他认为 DeepSeek 的巨大影响力，不仅在于其在 AI 领域实现的多项关键技术创新，更在于它引发了对全球 AI 生态格局升级和重塑的深刻思考。作为深耕开发者生态多年的行业大 V，李建忠盛赞 DeepSeek “或将成为中国开发者拥抱 AI 时代的最佳选择”。来自开发者社区的积极评价，无疑是最直接且最具说服力的市场反馈。

不仅如此，国际媒体也敏锐捕捉到了 DeepSeek 的崛起，并将其与具有历史意义的“人造卫星时刻 (Sputnik moment)”相提并论。“人造卫星时刻”源于 1957 年前苏联成功发射人类首颗

人造卫星 Sputnik 1 号，这一事件震惊世界，象征着科技竞争格局的突变，以及对原有技术领导者的挑战。外媒借用“人造卫星时刻”来形容 DeepSeek，一方面是认同 DeepSeek 在人工智能尤其是在大模型技术领域取得的突破性进展，另一方面是预示着其可能打破美国在 AI 领域的领先地位，并对全球 AI 竞争格局产生深远影响。这种来自国际舆论场的积极信号，进一步印证了 DeepSeek 在全球 AI 领域的影响力。

2. DeepSeek，真的是“国运级创新”吗？

但是，很多人可能会质疑，它凭啥，可以担当得起“国运级创新”这个评价？

DeepSeek 虽然开源，虽然它的性能优异媲美一流的商业大模型，同时成本低廉为顶尖商业大模型的几十分之一，而且拥有大量的技术创新，包括在算法上和工程上。但是这些优点在行业内并非独有。在它之前，Meta 公司的 Llama 系列、阿里公司的 QWen 系列、欧洲公司的 Mistral 系列等开源大模型，早已在业界积累了良好的口碑，且同样具备出色的性能和开源特性。近年来，开源大模型已经成为行业内的重要趋势，开源本身已经不再是稀缺资源。

而且技术特性上，让它成本最低的模型架构 MOE (Mixture of Expert) 也不是原创，Mistral 之前早就这么做了。而纯用深度学习来进行模型推理能力的增强，这也是 OpenAI 半年前发布 OpenAI O1 的做法，只不过它没有开源而已。所以有人质疑，即便 DeepSeek 在部分指标上超越了前辈项目，这种领先是否足以称为颠覆性突破？此外，DeepSeek 的模型虽然开源，论文虽然公开而且很详细，但底层技术路线的原创性究竟有多少？

它凭什么能够在短时间内获得如此高的评价，甚至被视为“国运级创新”？

笔者认为，DeepSeek 是堪称“国运级”的科技创新，不仅仅是它的性能好、成本低、开源这三个技术方面的特点，最重要的原因，是它初步达成了国内外的产业共识。而产业共识，尤其是中国人领导的产业共识，非常难得，非常有价值，之前基本没有，就是我顶它为“国运级创新”的理由。

关键是它走出了一条在保持跟国际顶尖大模型效果相当的同时，大幅度降低训练和部署成本的路径，同时又以开源的形式让广大开发者可以很方便的进行部署提供服务。它对国内外的产业界和学术界，做出了巨大的贡献。希望它能继续沿着这条路，追求 AGI 和 AI 普惠，同时开源开放，让大模型更快的应用到生产和生活的千行百业中去，为社会创造更大价值。

DeepSeek 与 Grok3 的对比

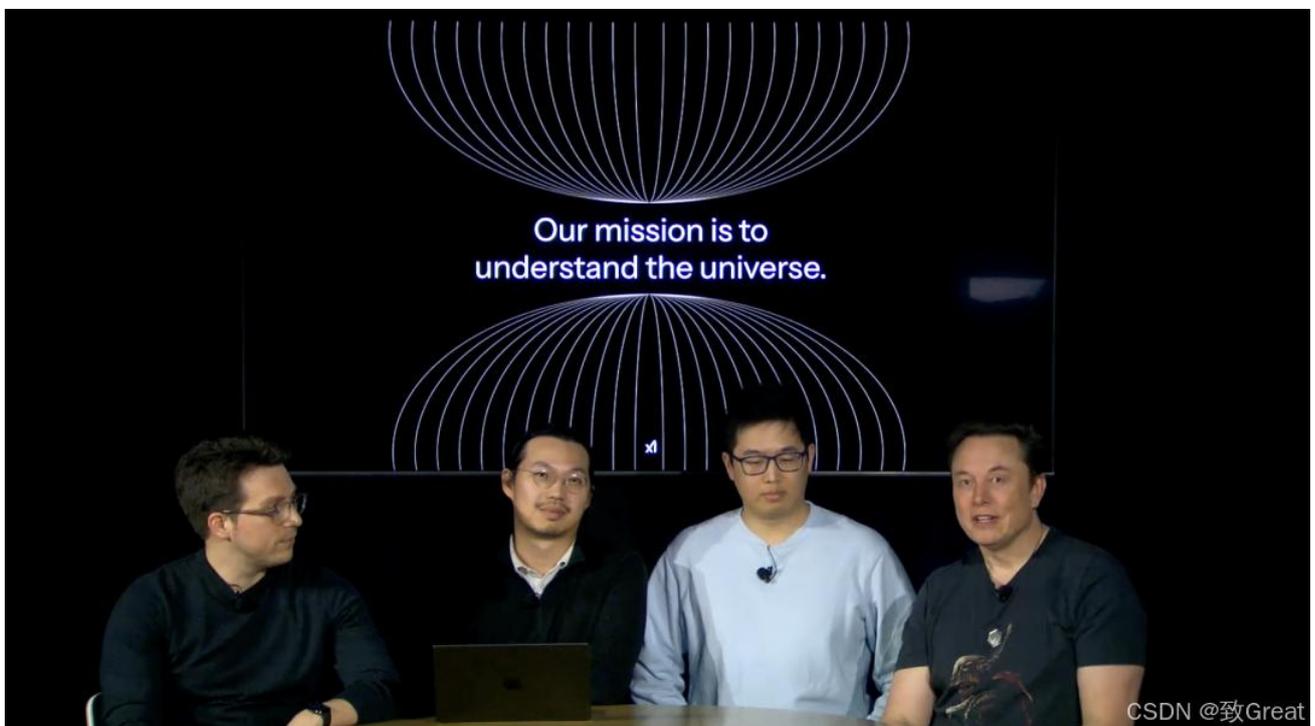
陈钟摘录

(从《关于 Grok3 和 DS 背后苦涩教训引发思考》一文中摘录)

引言

Ilya Sutskever (前 OpenAI 联合创始人兼首席科学家) 曾在在召开的 NeurIPS 会议上表示, 大模型的预训练已经走到了尽头。而 Noam Brown (OpenAI 研究员, 曾带领团队开发出在德州扑克中战胜职业选手的 AI 系统 Pluribus) 在关于 OpenAI O1 发布的采访中提到, **提升 Test-Time Compute 是提升大模型答案质量的关键**。2024 年的圣诞节前夕, 一片节日气氛下, 湾区的硅谷 AI 大佬、机构和投资者们正在深入探讨从 “Scaling Learning” 向 “Scaling Search” 转变的路径。而这一切的思考, 都可以追溯到 Rich Sutton (强化学习领域的奠基人之一) 在 2019 年发表的经典短文 The Bitter Lesson (苦涩的教训)。

从 2025 年初持续火爆的 DeepSeek-R1, 到近期马斯克带着团队全球直播发布了号称“史上最聪明 AI”的 Grok3, 一个是工程优化节省资源追求极致性价比, 一个是力大拍砖大力出奇无脑坚定地 **scale up**, 两条相似但又非常不同的路线, 恰如 AI 发展史上的两条岔路。这不仅是一场技术路线的对比, 更是对“苦涩教训”深刻内涵的再审视。



The Bitter Lesson: 苦涩教训

The Bitter Lesson（苦涩的教训）又称 AI 工程师的必背经典！



理查德·萨顿（Richard S. Sutton）：阿尔伯塔大学计算机科学教授，被认为是现代计算的强化学习创立者之一。他为该领域做出了许多重大贡献，包括：时间差分学习（temporal difference learning）、策略梯度方法（policy gradient methods）、Dyna 架构在内。

传说中 OpenAI 工程师需要背诵的经典文章《The Bitter Lesson（苦涩的教训）》，正是由他撰写。文章指出了过去 70 年来，AI 研究走过的最大弯路，就是过于重视人类既有经验和知识，而他认为最大的解决之道是摒弃人类在特定领域的知识、利用大规模算力的方法，从而获得最终胜利。

今天，就让我们静下心来，细读这篇雄文。或许在阅读之后，我们可以对当下与未来的 AI 发展，拥有更深刻的理解和启发，这篇文章信息量密度极高，重点都划不过来，强烈建议大家反复阅读！

以下中文内容我由 ChatGPT-4O 翻译+润色而来，欢迎大家阅读英文原文

原文作者：

The Bitter Lesson Rich Sutton March 13, 2019

原文地址：<http://www.incompleteideas.net/IncIdeas/BitterLesson.html>

从 70 年的人工智能研究中可以得出的最大教训是：利用算力的通用方法最终是最有效的，并且这种优势非常显著。其根本原因在于摩尔定律，或者更准确地说，就是单位算力成本随时间持续呈指数级下降这一客观规律。大多数人工智能研究都是在假设算力保持恒定的前提下进行

的（在这种情况下，利用人类知识几乎是唯一可行的提高性能的方法）。然而，若考虑比一般研究项目稍长的时间尺度，我们会发现，算力的增长是不可避免的，并且增长速度极快。在寻求短期内能带来提升的方法时，研究人员倾向于利用他们对特定领域的专业知识。然而，从长远来看，唯一重要的事情是如何利用算力。**领域知识和算力这两个事情客观来讲并不相关，但在实践中却常常相互冲突。**研究者的时间有限：**有限的时间投入到一种方法上，就必然会错失另一种方法带来的潜在收益。而且，研究者也容易对他们选择的方法产生心理上的依赖，甚至把这种依赖当做自己的特长，从而难以跳出固有思维。**更重要的是，基于人类知识的方法往往会增加研究的复杂性，限制了基于算力的通用方法的有效性。这条真理，是许多 AI 研究者在无数次的碰壁之后才痛彻心扉地领悟到的，其中一些典型的案例值得我们深思。

在计算机国际象棋领域，1997 年击败世界冠军卡斯帕罗夫的方法，主要依赖于大规模的深度搜索。当时，大多数研究计算机国际象棋的研究人员对此感到失望，因为他们所采用的方法是基于人类对国际象棋的特殊理解。当一个更简单的、基于搜索的方法，结合专门的硬件和软件，证明其效果远远优于人类知识驱动的方法时，这些研究人类知识的象棋研究者无法坦然接受失败。他们认为，“蛮力”搜索可能在这次比赛中取胜，但它不是一种通用策略，而且这也不是人类下棋的方式。这些研究者希望基于人类经验的方法能胜出，但现实让他们感到失望。

在计算机围棋领域，类似的研究进展模式再次出现，不过这一过程比国际象棋晚了 20 年。最初，大量的研究努力都集中在如何避免搜索，而是利用人类知识或围棋的特殊特性。然而，随着大规模搜索的有效应用，所有这些努力都变得无关紧要，甚至起到了反作用。另一个关键因素是自我对弈学习（self-play）的方法，它用于学习价值函数。这种方法在其他许多游戏中，包括国际象棋中都发挥了重要作用（尽管 1997 年击败世界冠军的国际象棋程序并没有依赖学习）。自我对弈学习，以及学习方法本身，与搜索类似，它们都能利用大量算力。搜索和学习是人工智能研究中利用算力的两大核心技术。在计算机围棋研究中，研究人员最初的努力集中在如何利用人类理解（从而减少搜索需求），但最终，通过拥抱搜索和学习，才取得了更大的突破。

在语音识别领域，1970 年代 DARPA 曾举办过一次早期竞赛。当时参赛的方法包括大量依赖人类知识的特殊技术，例如利用对单词、音素和人类声道结构的理解。而另一类方法则是更具统计性的方式，依赖大量计算，并基于隐马尔可夫模型（HMMs）。最终，统计方法战胜了基于人类知识的方法。这一结果引发了整个自然语言处理领域的重大变革，统计和计算方法逐渐在几十年间成为主导。最近，深度学习的兴起进一步推动了这一趋势。在语音识别领域，深度学习方法更少依赖人类知识，而是使用更多的算力，并在海量训练数据上进行学习，从而显著提高了语音识别系统的性能。和游戏领域一样，研究人员最初总是尝试构建符合他们对自身思

维过程理解的系统——他们希望将这些知识直接植入系统中——但事实证明这最终是适得其反的。而随着摩尔定律的发展，大量算力的可用性使得基于计算的方法取得了更好的效果，研究人员却为此浪费了大量时间。

在计算机视觉领域，类似的模式也曾上演。早期的计算机视觉方法主要基于寻找边缘、广义圆柱体，或使用 SIFT 特征等方法。然而，如今这些方法已基本被淘汰。现代深度学习神经网络仅依赖卷积（convolution）和某些不变性假设，最终取得了远超传统方法的性能。

这个教训非常深刻。然而，作为一个领域，我们仍然没有完全吸取这一教训，因为我们仍在重复类似的错误。要意识到这一点，并有效地避免，我们需要理解这些错误的吸引力。我们必须学会这个苦涩的教训：试图构建符合人类思维模式的人工智能系统，从长远来看是行不通的。这个教训建立在如下历史观察之上

1. AI 研究人员往往尝试将人类知识直接植入智能体；
2. 这种方法在短期内确实有所帮助，并且对研究者来说很有成就感；
3. 但从长远来看，它会停滞不前，甚至阻碍进一步的进步；
4. 重大突破最终是通过相反的方法实现的，即基于搜索和学习，利用大规模计算资源进行扩展。

最终的成功带有一种苦涩，因为它是对研究者所钟爱的“人类中心”方法的否定，这种成功往往难以被完全接受。

从这个苦涩的教训中，我们应该吸取的第一个重要经验是**通用方法的巨大威力**。只有那些能够持续利用增长算力的方法，才能在未来保持高效。目前来看，能够无限扩展的两类方法是**搜索和学习**。

第二个需要吸取的重要经验是：**人类思维的实际内容极其复杂，并且这种复杂性是无法简化的**。我们应该放弃试图以简单方式理解思维内容的做法，例如试图用简单方法来理解空间、物体、多智能体系统或对称性。这些概念都只是外部世界的一部分，它们本质上是复杂且无穷无尽的。我们不应该试图直接构建这些知识，因为它们复杂性是不可避免的。相反，我们应该构建能够发现和捕捉这些复杂性的元方法（meta-methods）。这些方法的关键在于它们能够找到良好的近似解，而我们应该让 AI 系统通过搜索和学习找到这些解，而不是由我们手动去定义它们。我们希望 AI 智能体具备像人类一样的发现能力，而不是仅仅包含我们所发现的知识。将人类已有的发现直接植入 AI 系统，反而会使得理解发现过程变得更加困难。

DeepSeek 与 Grok3 的对比

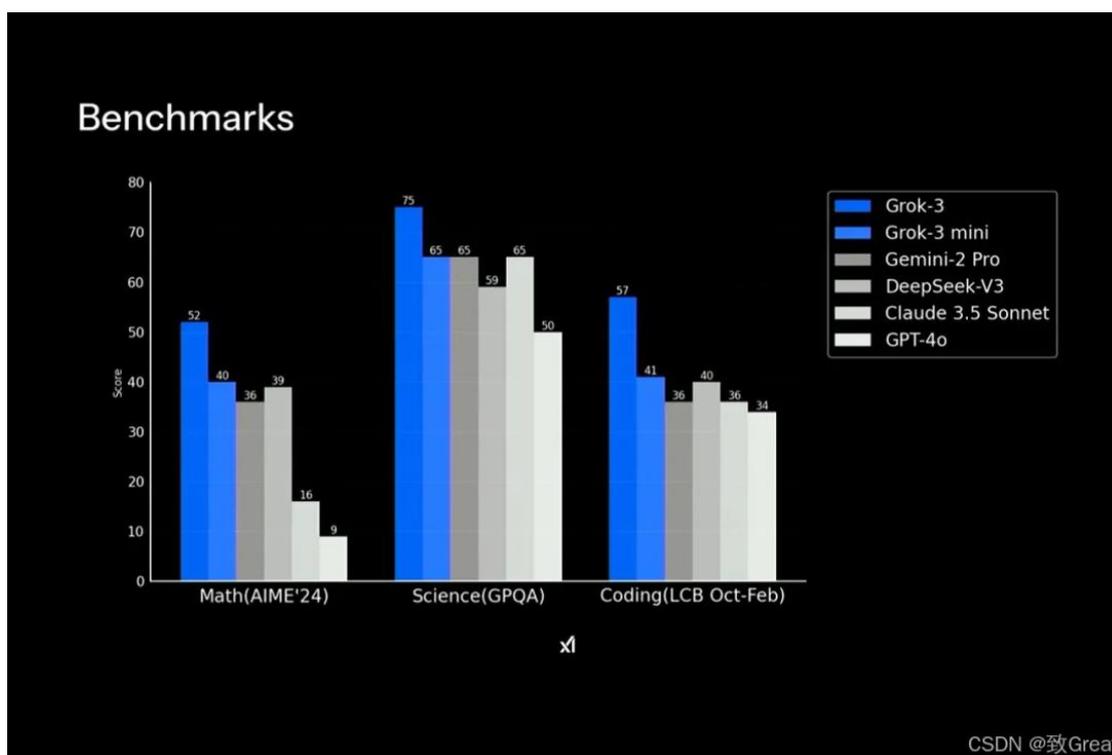
xAI 由埃隆·马斯克于 2023 年 3 月创立，目标是“理解宇宙的真正本质”。团队由来自 Google DeepMind、OpenAI 和 Microsoft 的顶尖专家组成，包括首席工程师 Igor Babuschkin 和顾问

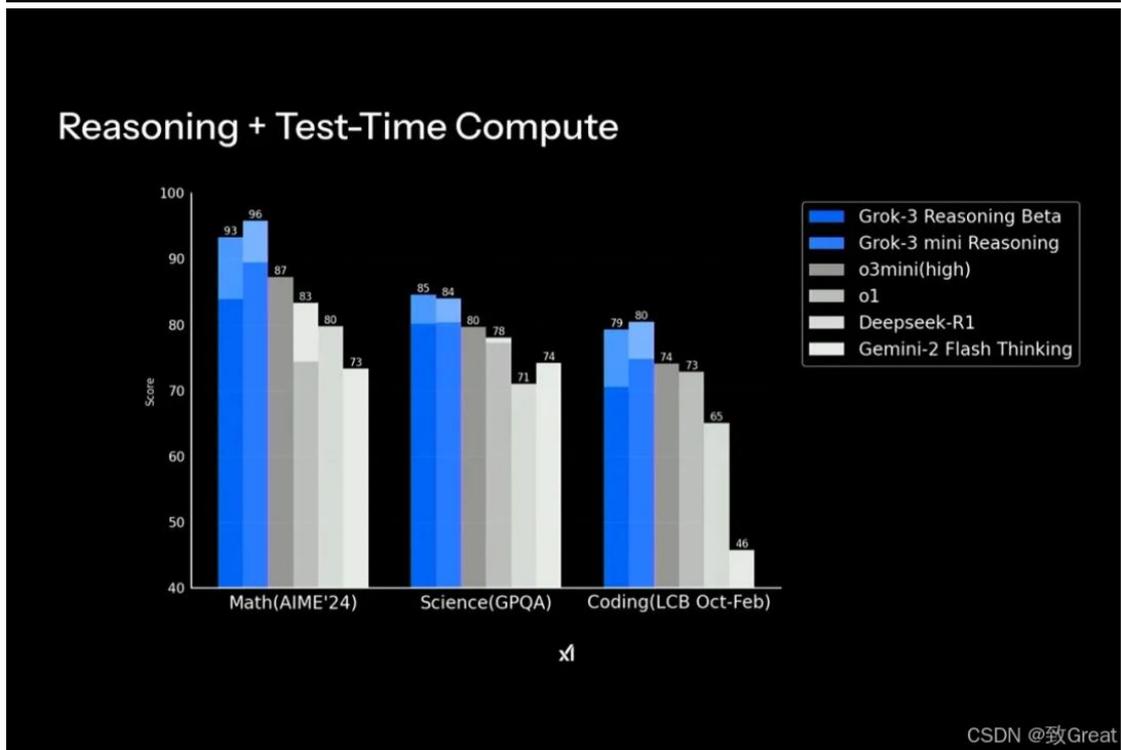
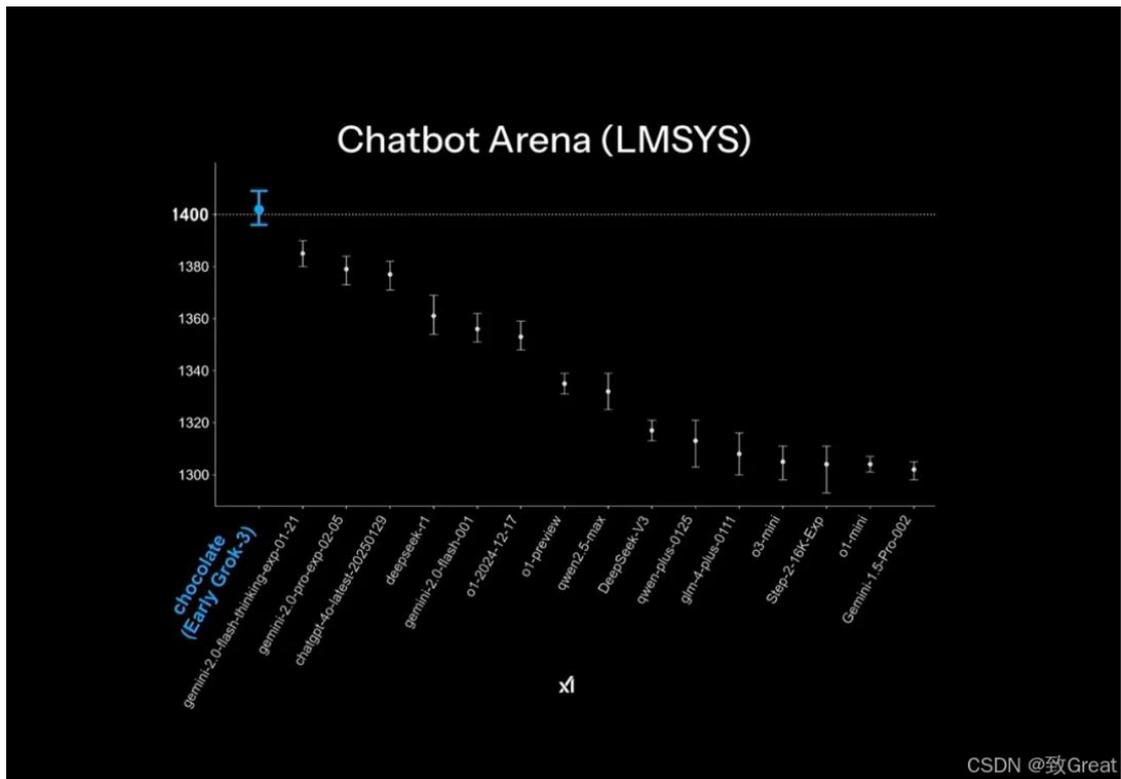
Dan Hendrycks。Grok 是 xAI 开发的 AI 聊天机器人，采用 Mixture of Experts (MoE) 架构，最新版本 Grok 3 于 2025 年 2 月发布，声称在多项基准测试中超越了 OpenAI 的 GPT-4o 以及其最新的推理模型 o1。

Grok 的发展历程体现了 AI 领域的快速迭代：

- Grok-1（2023 年 11 月）：3140 亿参数，64 层，8 个专家模型。
- Grok-1.5（2024 年 3 月）：提升推理能力，上下文长度达 128,000 tokens。
- Grok-2（2024 年 8 月）：进一步提升推理和多语言支持。
- Grok-3（2025 年 2 月）：使用 Colossus 超级计算机训练，配备 10 万至 20 万 NVIDIA H100 GPU。

Grok 的独特之处在于它的幽默感和通过 X 平台获取实时信息的能力，以及较少受限的回答风格。但更重要的是，Grok 3 的成功离不开 xAI 在算力上的巨大投入。xAI 在孟菲斯建造的 Colossus 超级计算机，配备了 10 万至 20 万 NVIDIA H100 GPU，为 Grok 3 的训练提供了强大的支持。





Grok 基座模型分别领先了 DeepSeek-v3 Gemini 和 GPT4o；推理模型在 benchmark 上得分也高于 OpenAI 的 o1/o3 系列和 deepseek-R1 模型。

Alberto Romero 在他的文章「Grok 3: Another Win For The Bitter Lesson」中写道：“Grok 3 的发布是「苦涩的教训」的又一次胜利”。xAI 没有像一些资源有限的公司那样需要极致的优化，

而是直接利用大规模的算力来训练模型。xAI 正是通过这种方式，快速追赶上了 AI 领域的领先者。

和国内 AI 公司背靠互联网大厂不同，DeepSeek 由量化对冲基金幻方量化运营。2016 年，幻方量化推出首个 AI 模型，第一份由深度学习生成的交易仓位上线执行。到 2017 年底，几乎所有的量化策略都采用 AI 模型计算。2023 年，幻方量化宣布成立创新性大模型公司 DeepSeek。

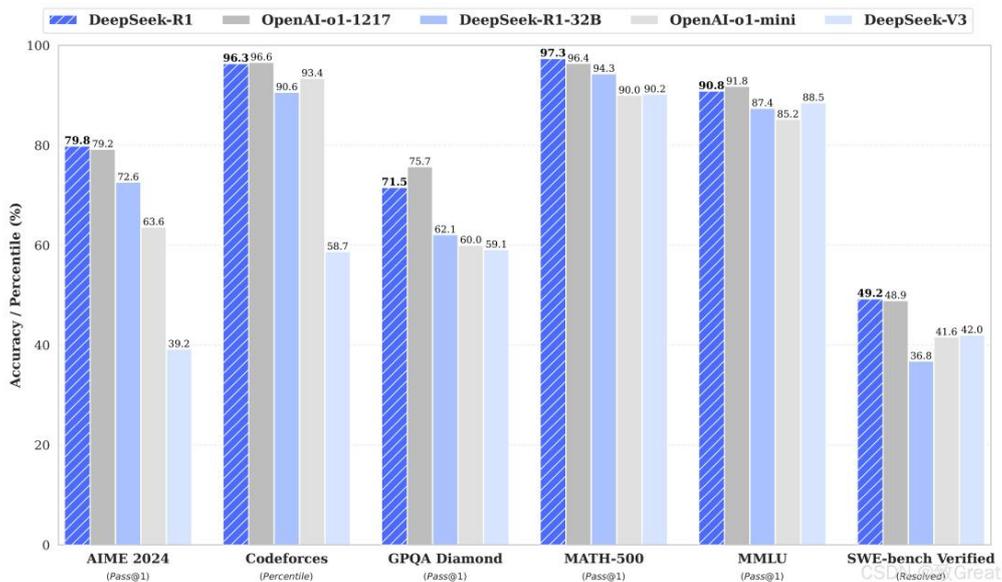
- 2024 年 12 月，DeepSeek-V3 首个版本上线，并同步开源，训练成本仅为 557.6 万美元，整个训练只需要 280 万个 GPU 小时。
- 2025 年 1 月 20 日，DeepSeek 又正式开源 R1 推理模型。1 月 24 日，DeepSeek-R1 在 Chatbot Arena 综合榜单上排名第三，与 OpenAI 的顶尖推理模型 o1 并列。

DeepSeek 之所以引发科技圈关注最主要的原因在于其以极小的成本训练出能和 OpenAI 的 ChatGPT 一较高下的人工智能大模型。DeepSeek 的 R1 的预训练(V3)费用只有 557.6 万美元，仅是 OpenAI GPT-4o 模型训练成本的不到十分之一。同时，DeepSeek 公布了 API 的定价，每百万输入 tokens 1 元（缓存命中）/4 元（缓存未命中），每百万输出 tokens 16 元。这个收费大约是 OpenAI o1 运行成本的三十分之一。

Training Costs	Pre-Training	Context Extension	Post-Training	Total
in H800 GPU Hours	2664K	119K	5K	2788K
in USD	\$5.328M	\$0.238M	\$0.01M	\$5.576M

Table 1 | Training costs of DeepSeek-V3, assuming the rental price of H800 is \$2 per GPU hour

年初 DeepSeek-R1 正式发布，并同步开源模型权重。DeepSeek-R1 基于 DeepSeek-V3 在后训练阶段大规模使用了强化学习技术，在仅有极少标注数据的情况下，极大提升了模型推理能力。在数学、代码、自然语言推理等任务上，性能比肩 OpenAI o1 正式版。



DeepSeek 与 Grok3 背后的苦涩教训

Grok3——规模的突破与代价

Grok3 是 xAI 在 2025 年推出的旗舰模型，被 Elon Musk 誉为“人类科学发现的超级加速器”。其训练过程堪称“暴力美学”的典范：**200,000** 个 Nvidia H100 GPU、数亿甚至十亿美元的投入，以及海量数据的预训练和微调。这种“大力出奇迹”的策略直接呼应了 Rich Sutton 在《苦涩教训》中提出的核心观点——**计算规模的增加最终会压倒一切人为设计的精巧方法**。

然而，这种成功的背后隐藏着苦涩的代价。首先，资源的消耗令人咋舌。200,000 个 H100 GPU 的训练不仅需要巨额资金，还伴随着巨大的能耗和环境影响。据估算，训练 Grok3 的电费可能相当于一个小型城市的年用电量。其次，规模化的收益似乎正在逼近天花板。尽管 Grok3 在推理能力和通用知识上表现出色，但与上一代模型相比，其边际提升是否匹配投入成本，仍是一个悬而未决的问题。

这种无脑堆积计算资源继续沿着 scaling law 曲线发展的路子，对咱们国内的很多公司来说，是可望而不可及的。

Grok3 的“苦涩教训”是：规模可以带来突破，但也可能成为一种依赖。当计算资源成为竞争的核心壁垒时，AI 开发的门槛被推向了只有巨头才能企及的高度。这种路径虽然短期内有效，却让中小型团队望尘莫及，同时也让人们开始反思：**如果“Scaling Learning”走到尽头，单纯的算力堆砌还能走多远？**

DeepSeek-R1——效率的创新与挑战

与 Grok3 的“土豪式”训练形成鲜明对比，DeepSeek-R1 代表了另一种哲学：**在有限资源下通过技术创新追求极致性能**。这款由中国初创公司 DeepSeek 推出的开源模型，建立在 DeepSeek-V3 基础模型之上，通过强化学习（RL）和监督微调（SFT）优化推理能力。其训练成本仅为约 600 万美元，相较于动辄数亿美元的对手，堪称“性价比之王”。

DeepSeek-R1 的成功得益于一系列高效训练技术。例如，它采用了基于规则的奖励系统替代传统的神经网络奖励模型，显著降低了计算需求；同时，通过“冷启动数据”和 RL 的结合，减少了对昂贵人工标注数据的依赖。这种方法不仅降低了成本，还通过开源模式推动了社区协作，使其影响力迅速扩大。

然而，这种创新并非没有代价（可能的猜测，类似我们要做实验试错）。DeepSeek-R1 的训练过程可能存在的**技术风险**：纯 RL 训练曾导致模型输出可读性下降和语言混杂问题，迫使团队引入额外的 SFT 步骤进行矫正。这种“试错式”开发暴露了效率路径的苦涩一面——**创新可能带来突破，但也伴随着更高的失败概率和调试成本**。其性能令人惊艳，在面对资源雄厚的对手时，DeepSeek-R1 的长期竞争力仍需时间验证。

DeepSeek-R1 的“苦涩教训”是：效率和创新可以在资源有限时打开局面，但技术复杂性和不确定性可能是其天花板。当规模化的对手继续加码时，效率之路是否能持续领先？

两条路线的碰撞与启示

Grok3 和 DeepSeek-R1 的对比，就像 AI 发展史上规模与效率、蛮力与智慧的永恒对决。Grok3 遵循《苦涩教训》的经典教条——计算规模是王道；而 DeepSeek-R1 则试图挑战这一教条，通过工程优化和算法创新突破资源瓶颈。两者的训练资源对比清晰地展示了这种差异，下面我们看看：

模型	训练资源 (GPU)	总成本 (估计)	方法论
Grok3	200,000 H100	数亿美元	预训练 + 微调
DeepSeek-R1	约 50,000 H100 估算的，可能不准	约 600 万美元	RL + SFT

苦涩教训的再审视

Rich Sutton 的《苦涩教训》曾指出，AI 的进步往往依赖于通用方法和计算能力，而非人类设计的特定规则。然而，DeepSeek-R1 的成功似乎在提醒我们：当计算资源成为稀缺品时，创新的“杠杆效应”可能比单纯的 **scale up** 更具价值。反过来，Grok3 的高成本则表明，规模化的道路虽然稳健，但可能陷入“资源诅咒”——过度依赖硬件而忽视算法层面的突破。

这种碰撞引发了更深层次的思考：

1. 从 **Scaling Learning** 到 **Scaling Search**

Ilya Sutskever 和 Noam Brown 的观点正在成为现实：预训练的红利逐渐耗尽，未来的突破可能在于推理时的计算优化 (Test-Time Compute)。DeepSeek-R1 的 RL 训练和 Grok3 的推理能力提升，都在不同程度上验证了这一趋势。

2. 资源的民主化与集中化

Grok3 的路径强化了巨头垄断，而 DeepSeek-R1 的开源模式则试图将 AI 技术民主化。两者孰优孰劣，或许取决于未来 AI 生态是倾向于集中式控制还是分布式创新。

3. 可持续性与伦理

Grok3 的高能耗引发了对 AI 环境影响的担忧，而 DeepSeek-R1 的低成本则提供了一种更可持续的范式。AI 的未来是否需要在性能与责任之间找到平衡？

DeepSeek 在显卡资源有限的情况下，把工程优化做到了极致，但如果有更多高性能 GPU，当然会更好，这点毫无疑问。换句话说，短期内通过工程上的优化来节省计算资源、提升性

能，并不违背计算资源的 scaling law，二者并不冲突。而且从长期来看，英伟达（Nvidia）的优势依然稳固。

今天看到网上不少人在讨论 Grok-3 用了这么多 GPU，但相比 DeepSeek 提升似乎没有预期中那么大。其实这个很好理解，因为 Grok-3 主要是和自家系列对比，来验证性能和计算量之间的关系——这就是典型的控制变量。

未来的趋势很清晰：在算力占据主导地位的前提下，不断优化工程效率，尽可能逼近算力规模所决定的模型极限。实际上，这次商用大模型的落地已经验证了 scaling law 依然有效。“The Bitter Lesson”也再次证明了一个道理——从长远看，算力始终是推动 AI 进步的决定性因素。从 Grok3 性能来看，尤其大家也不断尝试验证其 Deep Search 良好效果，当时 2025 年 1 月底英伟达股票暴跌时，当时市场可能有点过度悲观了。

结语：AI 未来的十字路口

Grok3 和 DeepSeek-R1 的故事，是 2025 年 AI 领域的一面镜子。Grok3 用规模证明了计算能力的无可替代，而 DeepSeek-R1 用创新展示了效率的潜力。然而，两者背后的“苦涩教训”告诉我们：没有一条路是完美的。规模带来胜利的同时埋下成本的隐患，效率带来希望的同时伴随挑战的阴影。

站在这个十字路口，AI 社区需要回答的是：未来的突破是继续“力大拍砖”，还是转向“精巧杠杆”？或许，正如《苦涩教训》的深刻启示，最终的答案不在于选择哪条路，而是如何在这两条路之间找到融合之道——用规模赋能创新，用创新优化规模。无论如何，Grok3 和 DeepSeek-R1 的较量，已经为我们打开了一扇思考未来 AI 发展的新窗口。

在年前举办的研讨会上，OpenAI 的核心成员 Noam Brown 发表了一系列重要观点。他强调尽管 Ilya 预训练模型的时代已经告一段落，但 Scaling Law 并未过时，而是进入了新的发展阶段。Noam 提出了 o1 模式，这种模式以推理计算为核心，优化算法和硬件设计，使 AI 在保持高性能的同时，大幅降低计算资源消耗。Noam Brown 认为，传统 Scaling Law 主要依赖于增加数据、计算资源和模型参数，而 o1 模式则是通过优化推理计算，提高模型效率。采用 o1 模式后，某些模型的推理速度提高 50%，而计算资源消耗减少 30%。o1 模式的提出不仅是对 Scaling Law 的革新，也是人工智能可持续发展的重要探索。通过优化推理计算，AI 未来可以更广泛应用于自然语言处理、计算机视觉、强化学习等领域，并推动技术普及。DeepSeek-R1 代表的高效计算路线，与 Grok3 代表的极致规模化路线，看似是两个完全不同的方向，但未来 AI 发展可能会融合这两种思路。我们可以预见以下几种可能的趋势：

1. **混合架构 (Hybrid Architecture)** : 未来的 AI 可能会结合 **专家模型 (Mixture of Experts, MoE)**、**DeepSearch (深度搜索)**，在需要的时候调用最合适的计算资源，而不是盲目扩大模型规模。
2. **动态计算 (Adaptive Compute)** : AI 可能会基于问题的复杂度调整计算资源，例如简单问题使用轻量推理，而复杂问题调用更强的搜索与规划能力。
3. **边缘计算 (Edge Compute)** : 未来 AI 的应用将更多部署到终端设备上，因此如何在有限的计算资源下优化 AI 推理将是关键。
4. **计算资源优化 (Compute-Efficient Training & Inference)** : 类似 DeepSeek 追求的高效计算方法，例如 **MLA、RL、FP8** 相关技术，将会成为未来大模型训练和推理的重要方向。

结合我们在做的事情，比如去年火爆的 RAG，从一开始简单 RAG、模块化 RAG 再到 Graph RAG, Agentic RAG，其中演化路径也是伴随着计算规模与效率优化的权衡利弊。在私有化部署以及应用场景中，受算力资源限制，可能应用了一个比较小参数量的大模型，其推理能力较差，在这种背景下我们需要做很多链路优化，比如用户问一个问题，可能经过查询改写、文档判断、信息压缩以及证据溯源等步骤。RAG 应用场景很广泛，模式很好，但是需要反复优化可能变得棘手繁琐，如今结合 DeepSeek-R1 以及 Grok3 的能力可以进一步将 RAG 优化变得更简单，也会出现更强的变体。

最终，未来 AI 竞赛的赢家将不仅仅是技术最先进的公司，而是那些能够在算力、技术创新和可持续性之间找到最佳平衡的公司。

点评 DeepSeekR1 与 Grok 3 Date 生成式 AI 排行榜比较

COPU

埃隆·马斯克(Elon Musk)最近发布了一款生成式 AI 语言大模型 Grok-3, 他对其有夸耀式的评价, 美国媒体还公布了包括 DeepSeeK-R1 和 Grok-3 在内的生成式 AI 推理模型的排行榜。本文拟先发表上述情况后再对其比较作一个点评。

马斯克对 Grok-3 最夸耀的评价是: Grok-3 将会是最后一个需要人类帮助的人工智能模型 (他的企业 XAI 有人说: GroK-3 让以前的模型都过时了)。

Grok-3 在标准测试中的跑分比 DeepSeek-R1 多了 10 分。Grok-3 的推理速度已达人类神经元的 1700 倍, 解题耗时仅为顶尖数学家的 1/23, 每 8.3 分可吞噬相当于维基百科的数据量, 模型权重 47 万次/分。当 Grok 的思维链开始自我进化时, 它将超越人类。

美国还公布了体现最近中美竞争的生成式 AI 排行榜

排行榜	推理模式	企业
1	Grok-3	XAI (美)
2	Gemini2.0 Flash Thinking	Google (美)
3	Gemini2.0 pro	Google (美)
4	ChatGPT 4o	OpenAI (美)
5	Deep Seek-R1	深度求索 (中)
6	Gemini 2.0 Flash	Google (美)
7	o1	OpenAI (美)
8	o1 Preview	OpenAI (美)
9	Qwen 2.5-Max	阿里巴巴 (中)

美媒还公布了 DeepSeek-R1 与 Grok-3 七项功能对比一览表

功能	DeepSeek -R1	Grok-3
1. 逻辑推理	结构清晰、较为机械	通俗易懂、易于理解
2. 编码与技术正确性	缺乏深度	提供详细、结构化、有洞察力的解释、优化细节
3. 现实世界知识与准确性	引用具体模型和技术显示其真实和近期发展	较为泛泛
4. 创造力	结构良好，略显平淡	其故事更具情感共鸣和动态性
5. 幽默感	笑话新颖、机智	笑话简单明了，常见
6. 辩论能力	结构清晰、缺乏动感性、说明力	提供更具吸引力，取得良好的回应，更贴近生活语言
7. 现实世界实用性	缺乏细节	提供一周餐食计划更实用

下面我们来作点评：

我认为对于生成式 AI 语言大模型 Grok-3 欲与 DeepSeek-R1 作比较，最不具可比性的是：为获得同等或相近的性能，DeepSeek-R1 的训练成本仅 600 万美元，而 Grok-3 等一些高端大模型的训练成本高达数万亿美元（即 DeepSeek-R1 训练成本仅为诸如 Grok-3 的 1/10 万~1/100 万），（而如我们以前讲过的 DeepSeek R1 的训练成本也已降至 GPT-4o 的 1/20）。

Grok-3 是在马斯克的豪华的训练服务中心（类似星际之门）进行训练的，该中心由 25 万张 H100 算力芯片卡集成建设而成的，综合算力可达 10^9 TFLOPS，需电力约 5000MW，巨额投资为 4 万亿美元。

问题是：Grok-3 训练需用 25 万巨量算力，投入 2 亿 GPU 小时，是颠覆性提升其性能吗？！DeepSeek 的横空出世宣告了美西方以巨额投资、高成本、堆叠资源，低改率，只供寡头发展 AI 的旧路的破产，而马斯克还想走这条旧路也是枉为其难。

因而对标比较 DeepSeek-R1 与 Grok-3 完全可一锤定音，DeepSeek R1 已拔得头筹！当然 DeepSeek 还有发展潜力和空间，假以时日，还可以进一步改善和提高性能。

从 DeepSeek R1 看大模型技术进展和趋势

北京智源人工智能研究院 倪贤豪

2024 H1 以来，大模型主流技术路线在往 Agent、推理等应用方向收敛。2024 年 11 月以来以 OpenAI o1/o3、Deepseek R1/R1 Zero 等为代表的推理模型密集发布即为此技术趋势的具体体现。

与此同时，预训练 Scaling Law 的技术布局和基础设施建设始终持续。尤其以海外头部厂商进展最为迅速，十万卡集群和下一代基础模型均已有所产出。

推理模型从技术难度和复现周期来看，难度相较下一代预训练基础模型而言并不大。技术难度来看，推理模型更多是基于以强化学习、MoE 等为代表的既有技术的系统组合“微创新”。而从复现周期来看，OpenAI o1 发布时隔 2 月，国内外推理模型便开始密集发布，也涌现出 Deepseek R1 等性能逼近 o1 而定价仅为其 1/30 的优秀复刻模型。相较于预训练 Scaling 而言，推理模型目前具有低成本、高性能提升的特点，这必定带来海量的新玩家尝试。低成本、高性能、海量玩家涌入，这三大特点或将加速推理模型的性能天花板在未来一年内快速见顶。

预训练 Scaling 在 GPT-4 以后，已有近两年没有大更新。从海外十万卡大集群建设布局加速、Grok3 和 Claude 3.5 Opus 等下一代基础模型相继完成训练来看，预训练 Scaling law 仍在生效，而下一代基础模型或将在未来半年到一年内陆续发布。相比于海外头部厂商十万卡以上集群建设如火如荼，国内头部互联网厂商十万卡集群建设将起步。近期全球非常火爆的 DeepSeek V3 模型，通过较少计算资源实现了逼近 GPT-O1 的效果，对国产模型有较强的现实激励意义。基于对 Deepseek V3 和 R1 技术报告的拆解，该系列模型所采取的技术创新多为对既有技术的模块替换和整合。“微创新”多过于变革性技术创新。一般而言，跟随者的路径更清晰，通常只用花 1/10-1/20 的成本即可跟随。但我们需要清醒地认知到，美国依然在投入巨大的资源进行算力建设和 AI 技术突破，下一代基础模型依旧会对模型全局性能带来提升。

一、近期国内外大模型进展

（一）明线：推理优化渐成大模型主流技术路线

自 2023 年以来，围绕基础模型的研究热度几经更迭，目前已从语言、多模态等基础模型，

往推理模型（以 o1 为代表）、Agent/RAG 等加速 AI 原生应用落地的方向迁移。

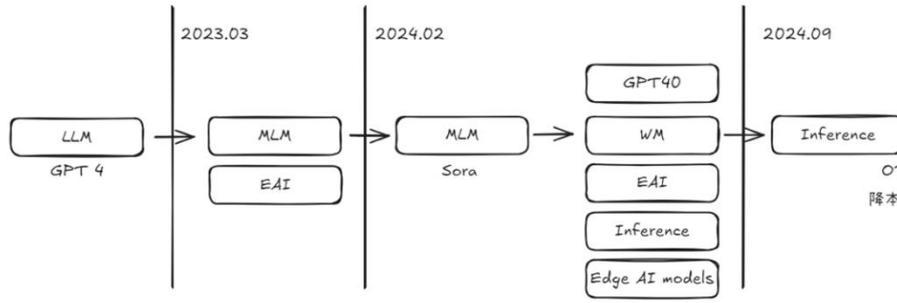


图 自 2023 年以来大模型技术路线收敛情况
数据来源：智源行研组

进入到 2024 年三季度，OpenAI 发布的 o1 模型带起推理模型的研究热潮，其在数学、编程等推理能力方面取得大幅提升，产学研研究重心也从预训练逐渐向后训练、推理阶段过渡。

o1 发布后，国内相关企业、高校及科研院所迅速跟进研发，并在约 2 个月时间内成功复现，自 2024 年 11 月起开始密集发布以复杂推理能力见长的推理大模型，且在多模态支持、代码生成、数理化问题解答等方向有不同侧重。

2024 年 11 月 16 日，月之暗面发布数学推理大模型 k0-math，该模型在解答数学问题时会深入思考和规划解题思路，在必要时进行自我反思和改进，以提高答题成功率；11 月 20 日，DeepSeek 发布 DeepSeek-R1，首发版本为 DeepSeek-R1-Lite-Preview，正式版则于 2025 年 1 月 20 日推出，并在开源社区引爆；2024 年 12 月 31 日，智谱 AI 上线推理大模型 GLM-Zero-Preview，即 GLM-Zero 预览版，擅长处理数理逻辑、代码和需要深度推理的复杂问题；1 月 15 日，科大讯飞发布星火深度推理模型 X1，星火 4.0 Turbo 底座全面升级，首发星火语音同传大模型；1 月 16 日，阶跃星辰发布推理模型 Step Reasoner mini，其既擅长通过超长推理能力，解决逻辑推理、代码和数学等复杂问题，亦能兼顾文学创作等通用领域。

与此同时，Agent、RAG 等其他应用技术也迎来密集更新。如 OpenAI 继发布 o1、o3 推理模型后，持续在 Agent 方向发力：更新 ChatGPT 的 Tasks 功能，并发布 AI Agent-Operator；MiniMax 开源基础语言模型 MiniMax-Text-01、视觉多模态模型 MiniMax-VL-01，支持 400 万长上下文，为 Agent 应用奠定基础；智谱发布 GLM-PC，目前已覆盖手机智能体 AutoGLM 和电脑智能体 GLM-PC 两大系统，实现了工具使用能力的深度突破。随着推理模型、Agent 等技

术持续迭代，AI 应用大规模落地的可能性在攀升。

（二）暗线：预训练模型、基础设施投入依然持续

尽管国内厂商在短期内难以跟进参数规模放大一到两个量级的基础模型，但也并未因此停止在预训练环节的算法创新。如 DeepSeek、MiniMax、Moonshot 等公司仍然围绕基础模型的训练算法、计算框架等方向优化，发布相关工作。

同时，大厂持续加码基础设施投入，筹备算力为依托的下一代基础模型。如字节跳动 2025 年计划投入超 120 亿美元建设人工智能基础设施，并启动“Seed Edge”研究项目，拟定 5 大研究方向，核心目标是做比预训练和大模型迭代更长期、更基础的 AGI 前沿研究。

值得一提的是，海外厂商在预训练和基础设施投入远快于国内，十万卡集群和下一代基础模型均已有所进展。如星际之门、AI 曼哈顿等项目或报告建议，甚至已开始有美国国家动员体制加速进入 AI 领域的迹象。

表 近期大模型领域重点进展情况

名称	发布时间	发布机构	聚焦方向
k0-math	2024.11.16	Moonshot	推理模型
DeepSeek-R1	2024.11.20	DeepSeek	推理模型
书生 InternThinker	2024.11.25	上海 AI Lab	推理模型
天工大模型 4.0 Skywork o1	2024.11.27	昆仑万维	推理模型
QwQ-32B-Preview（开源）	2024.11.28	阿里巴巴	推理模型
o1-Coder（开源）	2024.11.29	北京交通大学	推理模型
360gpt2-o1	2024.12.13	360	推理模型

名称	发布时间	发布机构	聚焦方向
k1	2024.12.16	Moonshot	推理模型
LLaVA-CoT（开源）	2024.12.17	北京大学、清华大学、鹏城实验室、达摩院、里海大学	推理模型
DeepSeek-V3（开源）	2024.12.30	DeepSeek	推理模型
QVQ-72B-Preview（开源）	2024.12.31	阿里巴巴	推理模型
GLM-Zero	2024.12.31	智谱 AI	推理模型
“日日新”融合大模型	2025.01.10	商汤科技	推理模型
MiniMax-Text-01、MiniMax-VL-01	2025.1.15	MiniMax	多模态模型；Agent（超长上下文，潜在应用于Agent）
X1	2025.1.15	科大讯飞	推理模型
Tasks 功能	2025.1.15	OpenAI	Agent
moonshot-v1-vision-preview	2025.1.15	Moonshot	多模态模型
Step Reasoner mini	2025.1.16	阶跃星辰	推理模型
k1.5	2025.1.20	Moonshot	推理模型
豆包大模型 1.5pro	2025.1.22	字节跳动	推理模型
Seed Edge 研究计划	2025.1.23	字节跳动	基础设施；AGI 路线

名称	发布时间	发布机构	聚焦方向
GLM-PC	2025.1.23	智谱 AI	Agent
Operator	2025.1.24	OpenAI	Agent
Qwen2.5 VL	2025.1.28	阿里巴巴	多模态模型
OpenAI o3-mini	2025.2.01	OpenAI	推理模型
OpenAI DeepResearch	2025.2.03	OpenAI	Agent
s1	2025.2.06	李飞飞	推理模型
Step-Video-T2V; Step-Audio	2025.2.18	阶跃星辰	多模态模型
Grok 3	2025.2.18	xAI	多模态模型

数据来源：智源行研组资料整理

二、推理上限很快到来，预训练投入仍需继续

（一）OpenAI 和 Deepseek 领跑 2025 年初推理模型赛道

以 OpenAI o1/o3 及 Deepseek R1 为代表的推理模型，引领了 2024 第三季度以来行业的主流技术趋势。

一方面，o1/o3 系列模型在推理模型的性能表现上依旧领跑整个赛道。

通过引入 RFT 及审慎对齐等关键技术，在推理过程中，利用思维链（CoT）分层解决复杂问题。o3 相较于 o1 在推理能力、编程性能、数学竞赛和科学问答准确率、安全性、复杂任务处理能力等多方面表现出明显的提升。

表 o1, o3 模型对比

特性	OpenAI o3	OpenAI o1
推理能力	显著增强，支持多步骤推理	基础推理能力
编程性能	SWE-Bench Verified: 71.7%	SWE-Bench Verified: 48.9%
数学竞赛准确率	AIME 2024: 96.7%	AIME 2024: 83.3%
科学问答准确率	GPQA Diamond: 87.7%	GPQA Diamond: 78%
Codeforces Elo	2727	1891
复杂任务处理能力	新任务上表现接近人类水平	对复杂任务处理能力有限
安全性	引入了“审慎对齐”机制，减少不安全输出	安全性较低，缺乏先进的对齐机制
适应性和灵活性	能根据任务复杂度调整推理深度	灵活性不足，适应性较差
计算效率	高计算模式下表现优异，能有效利用资源	效率较低，资源消耗较高

数据来源：OpenAI，智源行研组

o3 在 Epoch AI 的 FrontierMath 基准 25.2% 的得分（其他模型通常得分低于 2%），因近日被曝出提前拿到数据集，在此不做收录。但即便仅参考 AIME 2024 测试成绩，o3 的优异表现也已不言而喻。

另一方面，国内以 Deepseek R1 系列模型为代表，在性能逼近 o1 的前提下，实现训练成本和定价的远低于海外同行，加之模型本身的高度开源，推动 2025 年初的推理模型热度达到峰值。

o1 模型发布之后，国内厂商沿袭该路线快速推出推理模型，实现推理任务的能力提升，其中典型代表即为 Deepseek V3、R1 及 R1 Zero。其中，Deepseek V3 发布于 2024 年 12 月 26 日，R1 系列则发布于 2025 年 1 月 20 日。相较于 o1/o3 系列模型而言，根据官方介绍，DeepSeek-R1 在后训练阶段大规模使用了强化学习技术，在仅有极少标注数据的情况下，极大提升了模型推理能力。在数学、代码、自然语言推理等任务上，性能对标 OpenAI o1 正式版。

表 o1 与 DeepSeek R1 对比

比较维度		OpenAI o1-1217	Deepseek R1
数学能力	AIME 2024	79.20%	79.80%
科学问答	GPQA Diamond	75.70%	71.50%
编程性能	Codeforces (准确率)	96.6	96.3
	SWE-Bench Verified	48.90%	49.20%
自然语言能力	MMLU	91.80%	90.80%
API 价格 (美元/百万 Token)	输入	15	0.55
	输出	60	2.19

数据来源：Deepseek R1 技术报告

如上表，根据实际场景测试，目前 R1 仍难以与 o1 在所有场景对标，但在自然语言能力、数学能力、科学问答、编程性能等多个指标已十分接近。

考虑到两者在价格上的差异以及模型的高度开源，Deepseek R1 在商业应用和社区影响力上的潜力十分可观。

作为参考，我们以 API 定价为例，Deepseek R1 定价仅为 OpenAI o1 的 1/30。Deepseek R1 的输入 token 收取 0.55 美元/百万，输出 token 收取 2.19 美元/百万；而 OpenAI o1 对输入 token 收费 15 美元/百万，输出 token 收费 60 美元/百万。

可以说，2024 年末 2025 年初的这一波推理模型发布，因为 OpenAI o1/o3 的性能突破以及 Deepseek V3 和 R1 在性能逼近前提下的成本优势，而在业界达到了热度的峰值。

（二）推理模型技术门槛不高，复现时间较短

客观来说，在推理模型吸引全行业目光的当下，在技术层面，我们认为，推理模型的难度和资源成本并不大。当然，这一判断是基于和 GPT4、Claude 3.5 等基础模型的难度和资源成本的对比。

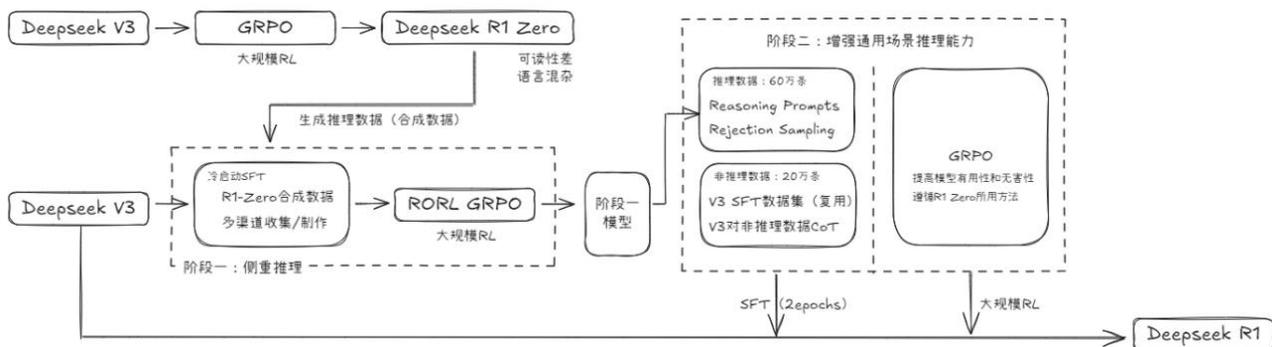


图 Deepseek R1 Zero 及 R1 技术原理图

数据来源：Deepseek R1 技术报告，智源行研组

1. Deepseek R1 Zero 和 R1 的技术实现

就 R1 的技术实现来看，Deepseek 团队主要基于 Deepseek V3 Base 基础模型，通过大规模强化学习（RL）及多阶段循环训练完成模型的训练收敛。

Deepseek R1 Zero: RL 技术储备（GRPO）和路线尝试

如上图，纯粹通过大规模 RL（GRPO）完成训练的 Deepseek R1 Zero，便已在 AIME2024 的 pass@1 score 实现从 15.6%到 71%的大幅提升。但由于该模型仍存在可读性差和语言混杂的问题，因此 Deepseek 团队将其用作生成推理数据集以支持 R1 模型训练。

值得一提的是，随着强化学习训练的推进，R1 Zero 的推理能力呈现出稳定且持续提升，且无需任何监督微调数据。通过利用更长的测试时间计算（Test-Time Compute），自然地获得解决越来越复杂推理任务的能力。这对于推理 Scaling Law 极为关键。

Deepseek R1 的多循环训练：

正式开始 R1 训练后，同样基于 Deepseek V3 base 模型，通过引入 R1 Zero 生成及其他渠道收集制作的推理数据集作为冷启动数据开始冷启动 SFT。

完成冷启动 SFT 后，通过采用和 R1 Zero 类似的定向推理 RL，完成第一个过程模型的训练收敛(阶段一模型)。通过冷启动 SFT 和定向推理 RL 进行训练，主要目的是为了解决 R1 Zero 的推理过程可读性差等问题。

随后，Deepseek 团队将阶段一模型再次用于生成推理数据集，与冷启动数据集专注于提

升模型可读性和定向推理能力不同，阶段一模型生成的推理数据结合多行业多场景，以增强模型在写作、角色扮演等通用场景的能力。除了阶段一模型生成的推理数据，同时合并其他数据扩展推理数据集，该部分扩展数据也采用 V3 模型做判断。该推理数据集总计约 60 万条。此外，Deepseek 团队另外复用了 V3 模型的 SFT 数据、V3 管道和其他精选的数据，也通过 V3 生成思维链，完成非推理数据集共 20 万条的构建。

基于以上共计 80 万条的推理和非推理数据集，重新采用 Deepseek V3 base，经过 2 个 epoch 的 SFT 及随后和 R1 Zero 类似的大规模 RL，收敛得到最终的 Deepseek R1 模型。

2. Deepseek V3 和 R1 系列模型的技术创新

基于对 Deepseek V3 和 R1 技术报告的拆解，该系列模型所采取的技术创新多为对既有技术的模块替换和整合。“微创新”多过于变革性技术创新。

我们围绕训练数据、模型架构、预训练、后训练及推理、系统架构等模型训推阶段和相应工具平台，对 Deepseek V3 和 R1 模型的创新加以说明。

训练数据方面，Deepseek R1 的训练采用了基于 R1 Zero 生成的长 COT 数据，经其他渠道数据扩展后构建高质量的冷启动数据集。采用冷启动数据 SFT 后，有效地对模型的规范化表达予以提升。

模型架构方面，Deepseek V3 采用了 DeepseekMoE 和 MLA Multi-head Latent Attention，（多头注意力）架构。

其中，DeepseekMoE 架构基于传统 MoE 架构改进（1991 年，Geoffrey Hinton & Michael I. Jordan）。相较专家数较少、调用均衡的传统 MoE，Deepseek 设计了每一层有 256 个路由专家、1 个共享专家。

同时，为避免专家数大量增加带来的辅助损失对模型性能的负面影响，Deepseek 团队采用了无辅助损失负载均衡（Auxiliary-Loss-Free Load Balancing），有效解决了该问题。

表 DeepSeek V3 与 R1 的创新技术对比

	V3	R1
模型架构	DeepseekMoE	-
	MLA	-
训练数据	-	冷启动数据、长 COT 数据
预训练	-	-
后训练	-	GRPO+多阶段循环训练
推理	-	-
系统架构	无辅助损失的负载均衡	-
	通信优化 (DualPipe)	-
	内存优化 (重计算、MTP)	-
	计算优化 (FP8)	-

注：R1 为基于 V3 多阶段训练模型，对于 V3 技术有继承关系，本表不做特别注释，R1 列仅显示 R1 专有创新技术。

数据来源：Deepseek V3 及 R1 技术报告，智源行研组整理

至于 MLA 架构，在 Deepseek V2 便已引入的该架构，通过将 Key(K)和 Value(V)联合映射至低维潜空间向量(cKV)，显著降低了 KV Cache 的大小，从而提升了长文本推理的效率。相较传统注意力机制，MLA 能将每次查询所需的 KV 缓存减少约 90%，极大的降低了推理成本。

系统架构方面，Deepseek 则采用了内存优化（重计算、MTP）、FP8 混合精度训练等策略，使模型训练更稳定，同时节省显存、加速训练。

后训练方面，Deepseek R1 通过 RL（强化学习）提升模型的推理能力。通过引入 GRPO（群组相对策略优化）算法，再基于奖励设计、训练模板等多阶段循环训练方式（基础->RL->SFT->RL->SFT->RL），实现通过大规模强化学习显著提升模型能力。

整体而言，通过对既有技术模块，尤其是强化学习（RL）的创新组合使用（长 COT 合成数据及多阶段训练），Deepseek R1 为代表的推理模型便实现了性能的显著提升。

3.关键技术是基于全球开源技术进行进一步迭代优化

除了业已提及的 MoE，Deepseek V3/R1 的多项关键技术，均是基于开源工作成果的再迭代优化。比如，MTP 方法即为 Meta 于 2024 年提出；FP8 混合精度训练，微软研究院在 2023 年 H100 发布即实现基于 FP8 混合精度训练，并发表了详细的技术文章。基于大规模强化学习，并结合多阶段训练的后训练流程，也是过去一年全球学术界和工业界广泛尝试、采用和优化的技术。

4.复现周期较短，也是对其目前技术难度较低的佐证

考虑到 o1 在 2024 年 9 月发布以来，时隔 2 个多月国内厂商便实现了相当程度的复现，自 2024 年 11 月起开始密集发布以复杂推理能力见长的推理大模型，且在多模态支持、代码生成、数理化问题解答等方向有不同侧重。

应该说，基于既有强化学习等既有技术的系统组合创新，从而实现模型性能提升，从技术难度和复现周期来看，后训练/推理 Scaling 难度相较于预训练 Scaling 更小，这对国内模型厂商而言是利好。当然，随着全球模型厂商的不断涌入，在技术和资源投入的持续深耕，预计在 2025H1 以后，推理模型的技术难度和成本也会不断增加，届时玩家数量将会收敛。

（三）后训练/推理 Scaling Law 天花板或将快速到来

推理模型的成本投入，远低于下一代基础模型所需的十万卡以上集群及其他相关资源需求。而其产生的性能提升仍相当可观，边际效应递减效应仍不明显。

以 OpenAI o1/o3、Deepseek R1 Zero/R1 为代表的推理模型，推动行业的范式变革。根据 Anthropic CEO Dario Amodei 的总结，我们可将该新范式表述如下：

基于预训练基础模型，再使用强化学习增加推理能力。

从目前的调研和技术拆解来看，o1/o3 强调 Test-Time Compute 为代表的推理 Scaling，R1 系列模型则更倾向于 GRPO+多阶段训练为代表的后训练 Scaling。值得注意的是，该范式对于预训练模型的能力要求并不低，前文提及的技术创新其实更多出自 V3 Base。

基于前文所提及的 OpenAI o3 及 Deepseek R1 引起的行业影响，我们在未来一年内将看到相当数量的大模型厂商尝试在以上方向有所突破。

Scaling Law类型	成本	性能	玩家数量	瓶颈时间
后训练/推理	+	++	+++	快速到来 (1年内?)
预训练	+++	+	+	仍有1~2代升级

图 Scaling Law 瓶颈期预测

数据来源：智源行研组

更低的成本投入、更高的性能提升，以及随之而来的更多玩家的涌入，或许将让我们很快看到推理 Scaling Law 天花板的到来。

（四）预训练 Scaling Law 仍在生效

在 2023 年 GPT4 发布以来，国内厂商囿于算力集群规模、高质量数据等原因，在预训练基础模型的追赶上备受挑战。随着全球 AI 主流技术路线往推理收敛，我们在 2024 年底开始看到了国内一众推理模型的发布。

与此同时，预训练 Scaling Law 并未失效。GPT-4 以来，尽管下一代预训练基础模型难产。但从目前实际调研信息来看，下一代基础模型 Grok 3 已发布，并在多个重要 benchmark 上体现出超越当前最优推理模型和上一代基础模型的能力；Claude 3.5 Opus 目前已接近训练完成或是在内部用以生成合成数据。而海外头部厂商在十万卡以上集群的搭建也正持续进行。

具体而言，作为新一代基础模型的首发代表，Grok 3 的能力提升依旧可观。在数学榜单 AIME2024 上，相较于 GPT 4o、Claude 3.5 Sonet 均值，Grok 3 提升达 316%，相较于 Gemini-2 Pro、Deepseek V3、Grok-3 mini 均值提升 38%；在科学榜单 GPQA 上，Grok3 相较于 Gemini-2 Pro、V3、Claude 3.5 Sonet、GPT 4o 均值提升 25%；在编程榜单 LCB 上，Grok 3 相较于其他主流模型均值提升达 56%。另外，Grok 3 整体数据集规模为 Grok 2 的 10 倍，其中数据合成引擎每天能生成约 50TB 的高质量训练数据，这表明提升算力、数据、模型规模从而提升模型表现的预训练 Scaling Law 仍然生效。

而另一方面，基于前文分析，推理模型的性能见顶在未来一年到来的概率不小。而随着后训练/推理 Scaling Law 开始呈现明确的边际效应递减，预训练 Scaling Law 的剩余价值将会出现比较优势。

为了更直观体现呈现比较优势后的预训练 Scaling Law 的价值。我们尝试围绕模型全局性能 (y) 和数据准备、训练、后训练、推理等关键变量做简单的建模说明：

$$y=f(x1,x2,x3,x4)+\Delta$$

其中，

x1: 数据准备质量（如数据清洗、标注准确性）

x2: 预训练过程（包括模型架构、超参数设置）

x3: 后训练过程（如微调、增量学习）

x4: 推理过程（如 Test-Time Compute）

Δ : 计算资源、优化策略、各阶段的交互影响等其他变量

可以认为，在大模型的生命周期中，如果某个环节（如数据准备、预训练、后训练或推理）提升了若干百分比，通常会对全局模型性能带来一定比例的提升。

以上述模型说明即为：后训练（x3）、推理（x4）等方向，经过未来一年的多方探索或将逼近性能天花板，同样的资源投入带来的性能提升将十分有限。

此时，仍有潜力可挖但成本巨大的预训练 Scaling Law 将体现出价值。预训练（x2）方向随着下一代基础模型发布，带来的性能提升，对全局模型性能的提升将越发凸显。一个典型佐证是 DeepSeek 团队在其 V2 模型上应用了与 V3 模型同样的多阶段循环训练方法，但效果不佳，可推断预训练模型性能对全局性能的影响举足轻重。

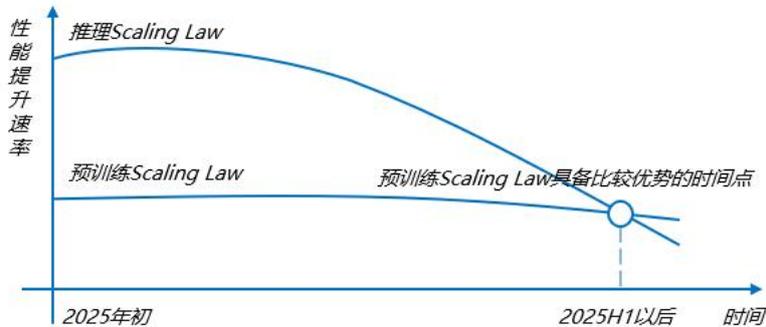


图 模型性能提升速率对比

数据来源：智源行研组

业界实践表明，大模型性能表现存在环节间的依赖性和性能传递效应。

前者表明大模型各个阶段是相互依赖的，一个环节的优化往往会影响到其他环节。

后者则意味着如果某一环节（如推理）已经达到了性能天花板，其他环节的提升仍然可以通过优化计算资源、改进模型架构或调整超参数等方式，进一步提高整体性能。这种性能传递效应意味着即使某一环节的改进效果有限，其他环节仍然能够发挥作用。

基于此，我们需要在模型性能提升与资源投入上有所权衡。不同环节不同阶段的性能提升所需资源不同，存在边际效应递减效应。

我们在模型训练中需要制定综合优化策略，实施资源分配的动态调整。

三、海外加速布局超大集群，新一代基础模型成果初现

根据目前国内外十万卡集群搭建进度及下一代基础模型训练进度来看，国内除了个别互联网头部厂商具备搭建十万卡集群的技术能力和资金储备，绝大多数大模型创业公司均已在该方向因算力集群出局。

至于海外，我们看到海外头部模型厂商在下一代基础模型和超大规模集群建设上已取得相当程度进展。

2024年7月，马斯克宣布包含有十万片 H100 GPU 集群 Colossus 正式投入使用，2025年2月发布的 **Grok 3 大模型**即在该超算中心完成训练；同时计划将在未来几个月内再增加 10 万片 GPU，其中 5 万颗将是更为先进的英伟达 H200，为下一代基础模型 Grok 4 的训练构建基础。

表 国外超大规模算力集群建设进展

公司	GPU	计划/建设规模	用途
Meta	H100	>100K	Llama4
xAI	H100/H200	100K（已完成）	Grok3/4
Anthropic	Trainium2	>100K（计划）	下一代 Claude

数据来源：Semianalysis，智源行研组

2024 年 11 月，Meta 首席执行官马克·扎克伯格在第三季度财报电话会议上透露，Llama 4 模型正在一个由十万片 H100 GPU 组成的集群上进行训练，同时，为了支持大模型的进一步训练，2025 年将进一步加大对 AI 基础设施的投资，总集群规模将达到 60 万片 H100 GPU。

AWS 在 2024 re:Invent 大会上宣布，AWS 与 Anthropic 共同构建一个名为 Project Rainier 的 EC2 UltraCluster 计算集群，该集群由 Trn2 UltraServers 组成，能够在数十万个 Trainium2 芯片之间通过第三代低延迟的 PB 级 EFA 网络进行分布式模型训练，其运算能力比 Anthropic 当前一代领先 AI 模型训练所需算力高出 5 倍以上。该项目将主要为 Anthropic 构建和部署未来模型使用。

根据 SemiAnalysis 信息披露，Anthropic 的新模型 Claude 3.5 Opus 已经训练完成，但因推理成本过高，商用性价比较低，因此目前主要用于生成训练数据及强化学习奖励建模。虽然商用的 Claude 3.5 Sonnet 为 Opus 的压缩版本，但其优秀的性能证明了基础模型继续 Scaling up 的潜力。

国内对于预训练仍需重视。

点评萨姆-奥特曼

COPU

2023.3.14, OpenAI 发布 GPT-4

奥特曼违反 2018.6 成立 OpenAI 时实行开源的初心, 采取闭源策略。

2024.5.3, OpenAI 发布 GPT-4o

2024.6.25, 奥特曼宣布对中国、俄罗斯、伊朗、朝鲜四国关停 GPT-4oAPI

2024.7.26, 奥特曼在《华盛顿邮报》上发表专栏文章, 呼吁建立以美国为首由西方世界主宰的 AI 主权联盟, 他反华反共, 反对、对抗中国侵占 AI 领域。

2025.1.31, 在《华盛顿日报》报导 DeepSeek 火爆出圈的当天, 奥特曼罕见地表态称, OpenAI 在开源 AI 软件方面, “一直站在历史的错误一边”。

美国科技媒体 VentureBeat 指出, 奥特曼最新表态承认了他的闭源策略可能存在缺陷, 随着来自中国的竞争加剧, 高效开放的开源模型越来越受欢迎, OpenAI 的闭源策略可能会发生重大转变。

奥特曼称赞 DeepSeekR1 是非常好的模型, 并将其研发的 GPT-4o 模型接入 DeepSeekR1 模型 B 端。

2025.1.31, 奥特曼发布首个开源免费的推理模型 GPT-3omini 系列 (low, medium, high), 并已上线运行。

2025.3.3, 奥特曼发布开源免费的 GPT-5。

2025.3.4, 奥特曼公开表示, 在发展人工智能方面愿与中国合作。

2025.3.5, 奥特曼邀请梁文锋访美谈 AI 合作, 后来他还公开表示, 每天醒来就担心被 DeepSeek 超越, 计划与 DeepSeek 的管理层交谈。

奥特曼又变脸了, 3 月 15 日他向国会提出要封杀 DeepSeek。

对奥特曼此人, 尚待听其言观其行!

阿里云千问 QwQ-32B 推理模型性能媲美 DeepSeek R1

宋逸群

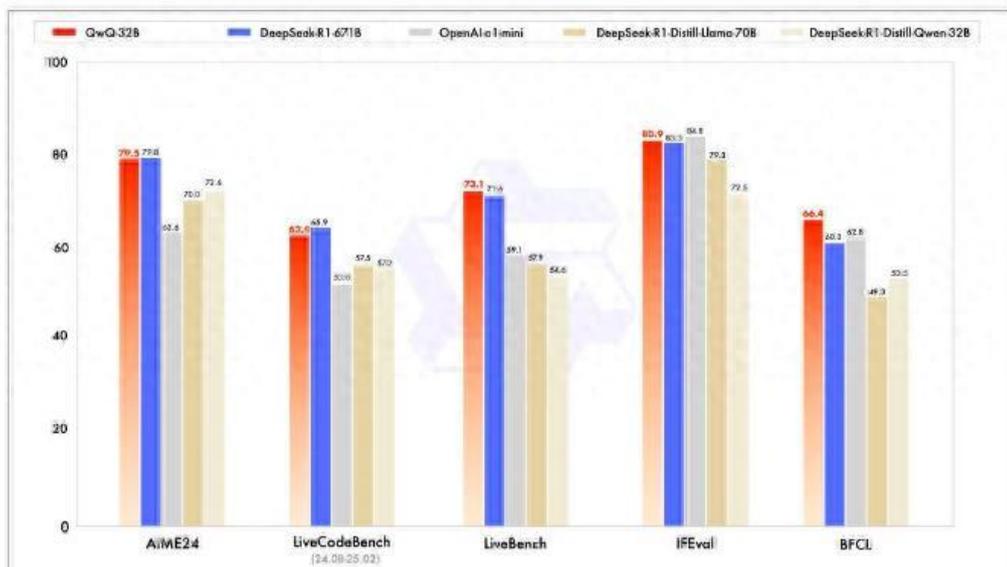
3月6日，阿里巴巴（09988.HK）跳空高开，一度涨至7.47%，报139.6港元/股，较2月24日的高点140.9港元/股，仅有一线之隔。

阿里云官方公众号发文称，公司发布并开源全新的推理模型通义千问 QwQ-32B。通过大规模强化学习，千问 QwQ-32B 在数学、代码及通用能力上实现质的飞跃，整体性能比肩 DeepSeek-R1。

在保持强劲性能的同时，千问 QwQ-32B 还大幅降低了部署使用成本，在消费级显卡上也能实现本地部署。此次，阿里云采用宽松的 Apache2.0 协议，将千问 QwQ-32B 模型向全球开源，所有人都可免费下载及商用。

阿里云指出，千问 QwQ-32B 性能比肩全球最强开源推理模型。

在一系列权威基准测试中，千问 QwQ-32B 模型表现异常出色，几乎完全超越了 OpenAI-o1-mini，比肩最强开源推理模型 DeepSeek-R1：在测试数学能力的 AIME24 评测集上，以及评估代码能力的 LiveCodeBench 中，千问 QwQ-32B 表现与 DeepSeek-R1 相当，远胜于 o1-mini 及相同尺寸的 R1 蒸馏模型；在由 Meta 首席科学家杨立昆领衔的“最难 LLMs 评测榜” LiveBench、谷歌等提出的指令遵循能力 IFEval 评测集、由加州大学伯克利分校等提出的评估准确调用函数或工具方面的 BFCL 测试中，千问 QwQ-32B 的得分均超越了 DeepSeek-R1。



据证券时报报道，业内人士分析，QwQ-32B 的成功表明，将强大的基础模型与大规模强化学习相结合，能够在较小的参数规模下实现卓越性能，这为未来通向通用人工智能提供了可行路径。

值得一提的是，据上海证券报报道，1月20日，DeepSeek-R1 推理模型发布后火速出圈，实现了最快复现 OpenAI o1 推理模型。业内预计，我国头部科技公司阿里巴巴、字节跳动等有望在 1 到 2 个月内跟上 R1 的创新成果。

1月29日，阿里发布通义千问 Qwen2.5-Max 大模型。在多个基准测试中，Qwen2.5-Max 比肩 Claude-3.5-Sonnet，并几乎全面超越了 GPT-4o、DeepSeek-V3 及 Llama-3.1-405B。

2月25日，通义千问官方账号 Qwen 在海外社交平台发布了 QwQ-Max-Preview（推理模型预览版），即将发布 QwQ-Max 正式版本、可部署在本地的较小版本以及官方 App QwQ-Max。

据了解，阿里通义团队一直坚持开源路线。从 2023 年至今，阿里通义团队已开源 200 多款模型，包含大语言模型千问 Qwen 及视觉生成模型万相 Wan 等两大基模系列，开源囊括文本生成模型、视觉理解和生成模型、语音理解和生成模型、文生图及视频模型等，覆盖从 0.5B 到 110B 等参数，实现了全模态、全尺寸大模型的开源。

长江证券研报指出，近期，阿里在 AI 领域持续发力，践行了其 AI 驱动战略，随着其后续投入的逐步提升，相关成果有望加速迭代，相关成果或将惠及相关产业链，加速 AI 应用落地，进而进一步带动算力需求的爆发。同时，随着阿里在 AI 基础设施、基础模型平台及 AI 原生应用、现有业务的 AI 转型等三方面加大投入，或将引领中国 AI 产业加速发展。

（声明：文章内容仅供参考，不构成投资建议。投资者据此操作，风险自担。）

再来谈谈我对 Manus 的见解

陆首群

2025.3.8

我感到网红 Manus 是一场炒作。它一开始就提出要与 DeepSeek 作对比，无非是想蹭 DeepSeek 的流量上马。我想谈谈 DeepSeek 的重大贡献，是找到或开创一条以低成本、高效率发展 AI 的新路，其重点不在于追求 AI 甚高的性能，实行开源也为其起到助力作用。

前几天我曾说过, Manus 主要是整合了外部很多模型的 APP。开始宣传 AI agent 虽然浮夸, 号称 Manus 是通用 AI agent 或真正自主智能体, 但重点还是放在通用上面, 说到自主智能还是离不开智能体, 但往后便不可收拾, 逐步滑向通用人工智能 AGI! 殊不知 AGI 的主要特点是具有高度自主的系统, 至今对它定义的内涵还不十分清晰。AGI 的出现将涉及 AI 是否挑战人类安全的问题, 以及人类应对之策的问题。说白了, 此时的 AI 其智能是否会超越人类? AI 是否摆脱人类并在其 AI 之间互相学习互相提高? 以后 AI 是否可能杀害或消灭人类? !如此重大严肃的问题岂是 Manus 轻飘飘一语带过, 所谓它已实现了 AGI, 但又无视于 AGI 对人类安全的挑战? !

还有一个小问题, 在 Manus 问世时, 网上对其捧得甚高, 谁在作点评? 我曾期望他们能提出点评的权威人士或测试数据, 但使我失望的是杳无音信, 看来除自媒体自吼外别无他人!

DeepSeek 提前发布 R2

5G 与 6G

3 月 11 日下午，有消息人士 DeepSeek 将于 3 月 17 日发布 DeepSeek-R2，5G 与 6G 公众号注意到较市场预期时间——今年 5 月份提前了近两个月。

1、技术突破方面

消息人士称 DeepSeek-R2 将有 3 大技术突破。

①**编程能力提升**。DeepSeek-R2 将具有更出色的编程能力，5G 与 6G 公众号认为可以让开发者借助它更高效地编写代码，缩短项目开发周期，提升软件的稳定性和功能性，能够在软件开发等领域提高生产效率，为相关企业节约人力成本。

②**多语言推理能力**。在全球化加速的当下，多语言推理能力的增强使 DeepSeek-R2 能理解和处理多种语言的复杂句式，为跨国企业、国际科研合作等提供支持，有助于其在全球市场获得更强的竞争力，满足不同语言背景用户的需求，促进信息流通和知识共享。

③**成本与准确性平衡**。DeepSeek-R2 以更低的成本提供更高的准确性，5G 与 6G 公众号认为意味着 DeepSeek-R2 不仅能够在性能上超越现有产品，还能通过优化算法或架构降低运营成本，能降低 AI 技术的应用门槛，让更多企业和机构能够负担得起，激发更多企业投入 AI 研发和应用，推动人工智能在更多领域的普及，形成良性循环。

2、市场竞争方面

提升行业地位。若上述特性能够兑现，DeepSeek-R2 可能会在全球 AI 竞赛中占据显著优势，使其成为全球 AI 市场中的佼佼者，进一步巩固 DeepSeek 的市场地位，重新定义企业战略和市场定位。

加剧行业竞争。在全球 AI 竞争激烈的当下，DeepSeek-R2 的发布将使行业竞争进入新阶段，给其他 AI 企业带来新的压力，促使它们加快技术研发和创新的步伐。

3、投资影响方面

带来重估机会。对于资产价格而言，DeepSeek-R2 的发布可能是一次重估机会。5G 与 6G 公众号认为相关行业如软件开发、翻译服务、数据安全等领域可能会因 R2 的应用需求提升，使得相关上市公司股价迎来估值提升。

投资风险并存。尽管市场对 DeepSeek-R2 普遍持乐观态度，但 AI 技术变化快速，存在技术被超越的风险，加上市场情绪可能过热，投资者需要保持理性，综合考虑各种因素，谨慎评估投资风险。



敬请关注联盟微信公众号
COPU开源联盟



扫描二维码
获取往期资料

中国开源软件推进联盟秘书处

电话：+86 010-88558999

联盟公共邮箱：office@copu.org.cn

联盟官网：<http://www.copu.org.cn>

地址：北京市海淀区紫竹院路66号赛迪大厦18层
