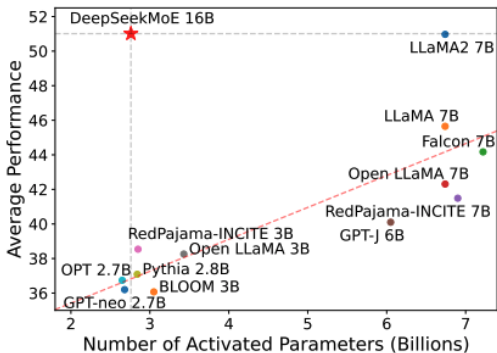
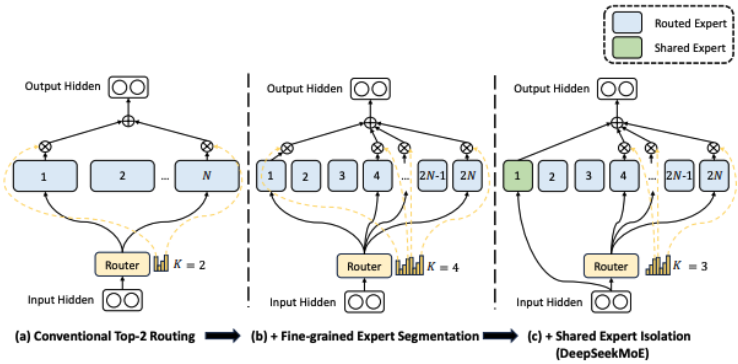


deepseek理论学习记录

模型名称	发布时间	参数量	主要特色																																										
DeepSeek LLM	2024.01.05	67B/7B	<p>DeepSeek LLM 7B/67B Base</p> <p>DeepSeek LLM 7B/67B Chat</p> <p>在2万亿token数据集训练，优秀的中英文能力，超越Llama2 70B</p> <p>通过SFT和DPO对基础模型进行优化，开发出Chat版本</p> <p>https://arxiv.org/pdf/2401.02954</p>																																										
DeepSeekMoE	2024.01.11		<p>1. DeepSeekMoE是一种新型专家混合架构，通过两个关键策略提升专家特化程度：将专家细分为mN个并从中激活mK个，以实现更灵活的专家组合；同时分离Ks个共享专家来捕获通用知识，减少路由专家的冗余。</p> <p>2. 在2B参数规模下，DeepSeekMoE就展现出与GShard 2.9B相当的性能，后者需要1.5倍的专家参数和计算量。DeepSeekMoE 2B的表现也接近具有相同参数量的密集模型。</p> <p>3. 扩展到16B参数时，DeepSeekMoE仅用约40%的计算量就达到了LLaMA2 7B的性能水平。进一步扩展到145B参数规模时，仅用28.5%的计算量就达到了与DeepSeek 67B相当的性能，显著优于传统GShard架构。</p> <div><table><caption>Data points estimated from Figure 1</caption><tr><th>Model</th><th>Number of Activated Parameters (Billions)</th><th>Average Performance</th></tr><tr><td>DeepSeekMoE 16B</td><td>~2.7</td><td>~51.5</td></tr><tr><td>LLaMA2 7B</td><td>~7.0</td><td>~50.5</td></tr><tr><td>LLaMA 7B</td><td>~7.0</td><td>~46.5</td></tr><tr><td>Falcon 7B</td><td>~7.0</td><td>~45.5</td></tr><tr><td>Open LLaMA 7B</td><td>~7.0</td><td>~44.5</td></tr><tr><td>RedPajama-INCITE 7B</td><td>~7.0</td><td>~43.5</td></tr><tr><td>GPT-J 6B</td><td>~6.0</td><td>~42.5</td></tr><tr><td>RedPajama-INCITE 3B</td><td>~4.0</td><td>~40.5</td></tr><tr><td>Open LLaMA 3B</td><td>~3.5</td><td>~39.5</td></tr><tr><td>Pythia 2.8B</td><td>~3.0</td><td>~38.5</td></tr><tr><td>BLOOM 3B</td><td>~3.0</td><td>~37.5</td></tr><tr><td>GPT-neo 2.7B</td><td>~2.7</td><td>~36.5</td></tr><tr><td>OPT 2.7B</td><td>~2.7</td><td>~36.0</td></tr></table></div> <p>Figure 1 Comparison between DeepSeekMoE 16B and open source models on the Open LLM Leaderboard. The red dashed line is linearly fitted from data points of all models except DeepSeekMoE 16B. DeepSeekMoE 16B consistently outperforms models with a similar number of activated parameters by a large margin, and achieves comparable performance with LLaMA2 7B, which has approximately 2.5 times the activated parameters.</p> <p>1. 图中用红色虚线表示除DeepSeekMoE 16B外所有模型的线性拟合趋势线。</p>	Model	Number of Activated Parameters (Billions)	Average Performance	DeepSeekMoE 16B	~2.7	~51.5	LLaMA2 7B	~7.0	~50.5	LLaMA 7B	~7.0	~46.5	Falcon 7B	~7.0	~45.5	Open LLaMA 7B	~7.0	~44.5	RedPajama-INCITE 7B	~7.0	~43.5	GPT-J 6B	~6.0	~42.5	RedPajama-INCITE 3B	~4.0	~40.5	Open LLaMA 3B	~3.5	~39.5	Pythia 2.8B	~3.0	~38.5	BLOOM 3B	~3.0	~37.5	GPT-neo 2.7B	~2.7	~36.5	OPT 2.7B	~2.7	~36.0
Model	Number of Activated Parameters (Billions)	Average Performance																																											
DeepSeekMoE 16B	~2.7	~51.5																																											
LLaMA2 7B	~7.0	~50.5																																											
LLaMA 7B	~7.0	~46.5																																											
Falcon 7B	~7.0	~45.5																																											
Open LLaMA 7B	~7.0	~44.5																																											
RedPajama-INCITE 7B	~7.0	~43.5																																											
GPT-J 6B	~6.0	~42.5																																											
RedPajama-INCITE 3B	~4.0	~40.5																																											
Open LLaMA 3B	~3.5	~39.5																																											
Pythia 2.8B	~3.0	~38.5																																											
BLOOM 3B	~3.0	~37.5																																											
GPT-neo 2.7B	~2.7	~36.5																																											
OPT 2.7B	~2.7	~36.0																																											

- 2. DeepSeekMoE 16B显著优于具有相似激活参数数量的其他模型，表现远超趋势线预期。
- 3. 尽管LLaMA2 7B的激活参数数量是DeepSeekMoE 16B的约2.5倍，但DeepSeekMoE 16B仍能达到与之相当的性能水平，显示出其架构的高效性。



- 1. 子图(a)展示了传统的top-2路由策略的MoE层结构。
- 2. 子图(b)展示了细粒度专家分割策略的实现方式。
- 3. 子图(c)展示了加入共享专家隔离策略后的完整DeepSeekMoE架构。

重要的是，这三种架构设计中专家参数数量和计算成本保持不变，说明DeepSeekMoE在不增加资源消耗的情况下优化了模型结构。

DeepSeek-Coder

2024.01.26

1B、5.7B、6.7B、33B四种规格

- 1. DeepSeek-Coder是参数规模1.3B到33B的开源代码模型系列，在2万亿tokens上训练，解决了代码智能领域过度依赖闭源模型的问题。
- 2. 在高质量项目级代码语料上预训练，采用16K窗口的填空任务来增强代码生成和补全能力，使用宽松许可证允许研究和商业使用。
- 3. 评估显示其性能超越了现有开源代码模型，也优于Codex和GPT-3.5等闭源模型，展现出卓越的代码生成和理解能力。

<https://arxiv.org/pdf/2401.14196>

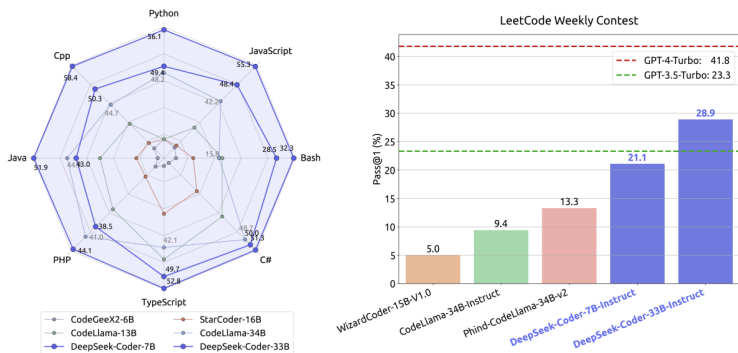


Figure 1 | The Performance of DeepSeek-Coder

DeepSeekMath	2024.02.05	基于以 DeepSeek-Coder-v1.5 7B为基础	<div><div><div>1. DeepSeekMath 7B是基于DeepSeek-Coder-Base-v1.5 7B模型，使用1200亿数学相关token继续预训练，数据来源包括Common Crawl、自然语言和代码数据。</div><div>2. 在竞赛级MATH基准测试中取得51.7%的成绩（使用64个样本的自一致性可达60.9%），接近Gemini-Ultra和GPT-4的水平，且无需依赖外部工具和投票技术。</div><div>3. 模型成功的关键在于精心设计的数据选择流程和新提出的群组相对策略优化（GRPO）算法，后者是PPO的变体，在提升数学推理能力的同时优化了内存使用。</div></div><div><table><caption>Figure 1 Top1 accuracy of open-source models on the competition-level MATH benchmark</caption><tr><th>Model</th><th>Approx. MATH Top@1 Accuracy (%)</th></tr><tr><td>LLaMA1-65B</td><td>10</td></tr><tr><td>WizardMath-70B</td><td>22</td></tr><tr><td>Qwen-14B</td><td>25</td></tr><tr><td>Mistral-7B</td><td>28</td></tr><tr><td>Llemma-34B</td><td>32</td></tr><tr><td>Qwen-72B</td><td>38</td></tr><tr><td>DeepSeekMath-7B</td><td>51.7</td></tr><tr><td>Gemini-Ultra / GPT-4 API</td><td>~50</td></tr></table></div><div><p>Figure 1 Top1 accuracy of open-source models on the competition-level MATH benchmark (Hendrycks et al., 2021) without the use of external toolkits and voting techniques.</p><p>https://arxiv.org/pdf/2402.03300</p></div></div>	Model	Approx. MATH Top@1 Accuracy (%)	LLaMA1-65B	10	WizardMath-70B	22	Qwen-14B	25	Mistral-7B	28	Llemma-34B	32	Qwen-72B	38	DeepSeekMath-7B	51.7	Gemini-Ultra / GPT-4 API	~50
Model	Approx. MATH Top@1 Accuracy (%)																				
LLaMA1-65B	10																				
WizardMath-70B	22																				
Qwen-14B	25																				
Mistral-7B	28																				
Llemma-34B	32																				
Qwen-72B	38																				
DeepSeekMath-7B	51.7																				
Gemini-Ultra / GPT-4 API	~50																				
DeepSeek-VL	2024.05.11	1.3B/7B	<div><div><div>1. DeepSeek-VL是一个开源视觉语言模型，数据构建方面覆盖网页截图、PDF、OCR、图表和知识内容等真实场景，并基于用户实际场景创建指令数据集进行微调。</div><div>2. 模型架构采用混合视觉编码器，可在固定token预算内高效处理1024x1024高分辨率图像，同时保持较低计算开销，以满足实际应用需求。</div><div>3. 训练策略强调语言能力的重要性，从以文本为主开始，逐步调整比例实现视觉和语言模态的平衡整合。DeepSeek-VL（1.3B和7B版本）在实际应用中展现出优秀的用户体验，并在多个视觉语言基准测试中达到同等规模最优或具有竞争力的表现。</div></div><div><p>https://arxiv.org/pdf/2403.05525</p></div></div>																		

Figure 1 | DeepSeek-VL possesses general multimodal understanding capabilities, capable of processing logical diagrams, web pages, formula recognition, scientific literature, natural images, and embodied intelligence in complex scenarios.

<https://arxiv.org/pdf/2405.04434>

<https://arxiv.org/pdf/2405.04434>

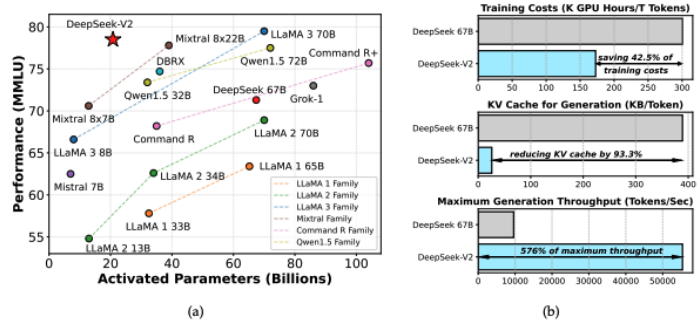


Figure 1 | (a) MMLU accuracy vs. activated parameters, among different open-source models. (b) Training costs and inference efficiency of DeepSeek 67B (Dense) and DeepSeek-V2.

DeepSeek-Coder-V2

2024.06.17

236B

1. DeepSeek-Coder-V2是一个开源的专家混合(MoE)代码语言模型，基于DeepSeek-V2的中间检查点，通过额外6万亿tokens的预训练显著提升了编程和数学推理能力，同时保持了通用语言任务的性能水平。
2. 相比DeepSeek-Coder-33B有显著进步：支持的编程语言从86种扩展到338种，上下文长度从16K扩展到128K。
3. 在标准基准测试中，DeepSeek-Coder-V2在编程和数学任务上的表现超越了GPT4-Turbo、Claude 3 Opus和Gemini 1.5 Pro等闭源模型。

<https://arxiv.org/pdf/2406.11931>

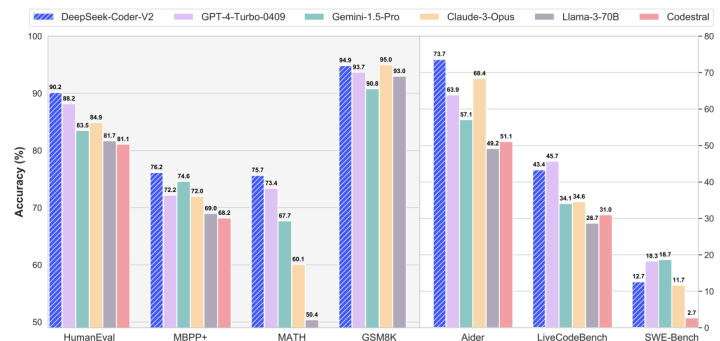


Figure 1 | The Performance of DeepSeek-Coder-V2 on math and code benchmarks.

DeepSeek-V2.5

2024.12.10

236B

结合DeepSeek-V2-Chat和DeepSeek-Coder-V2-Instruct的功能。这个新模型整合了前两个版本的通用能力和编程能力

Metric	DeepSeek-V2-0628	DeepSeek-Coder-V2-0724	DeepSeek-V2.5
AlpacaEval 2.0	46.6	44.5	50.5
ArenaHard	68.3	66.3	76.2
AlignBench	7.88	7.91	8.04
MT-Bench	8.85	8.91	9.02
HumanEval python	84.5	87.2	89
HumanEval Multi	73.8	74.8	73.8
LiveCodeBench(01-09)	36.6	39.7	41.8
Aider	69.9	72.9	72.2
SWE-verified	N/A	19	16.8
DS-FIM-Eval	N/A	73.2	78.3
DS-Arena-Code	N/A	49.5	63.1

DeepSeek-VL2

2024.12.13

10B-45B激活参数

- DeepSeek-VL2是一系列基于专家混合(MoE)的视觉语言模型，通过两大升级改进了前代产品：动态分块的视觉编码策略处理高分辨率图像，以及利用多头潜在注意力机制压缩Key-Value缓存实现高效推理。
- 在改进的视觉语言数据集上训练，在视觉问答、文字识别、文档/表格/图表理解和视觉定位等多个任务上展现出优秀性能。
- 提供三个版本：Tiny(1.0B)、Small(2.8B)和标准版(4.5B)激活参数，与现有开源密集模型和MoE模型相比，在相同或更少的激活参数下达到了竞争力或最优性能。

https://github.com/deepseek-ai/DeepSeek-VL2/blob/main/DeepSeek_VL2_paper.pdf

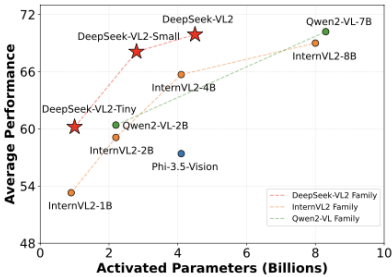


Figure 1 | Average performance vs. activated parameters among different open-source models. We average the accuracy of MMBench v1.1, MMStar, MMMU (Val), MathVista (TestMini), AI2D (Test), and OCRBench. The scores of OCRBench are divided by 10 to scale them to [0, 100].

DeepSeek-VL2采用llava风格的架构，由三个主要组件构成：视觉编码器(vision encoder)、视觉语言适配器(VL adaptor)和基于MoE的大语言模型。

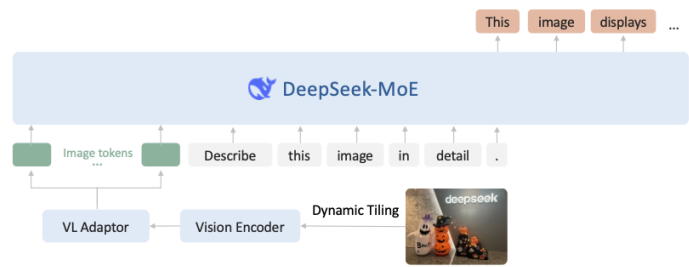


Figure 2 | Overview of DeepSeek-VL2. The overall structure is a llava-style architecture, which includes a vision encoder, a VL adaptor, and a MoE-based LLM.



Figure 3 | Illustration of dynamic tiling strategy in DeepSeek-VL2. By dividing images into multiple tiles, DeepSeek-VL2 achieves stronger fine-grained understanding capabilities compared to DeepSeek-VL.

1. DeepSeek-VL2采用动态分块策略，将输入图像划分为多个小块(tiles)进行处理，通过分块处理策略，模型能够更细致地关注图像的局部细节，增强了模型对细节的理解能力，相比DeepSeek-VL有明显提升。
2. 这种动态分块的方法特别适合处理高分辨率和不同纵横比的图像，使模型能更好地捕捉图像中的细微特征和空间关系。

DeepSeek-V3

2024.12.26

671B(激活37B)

架构创新

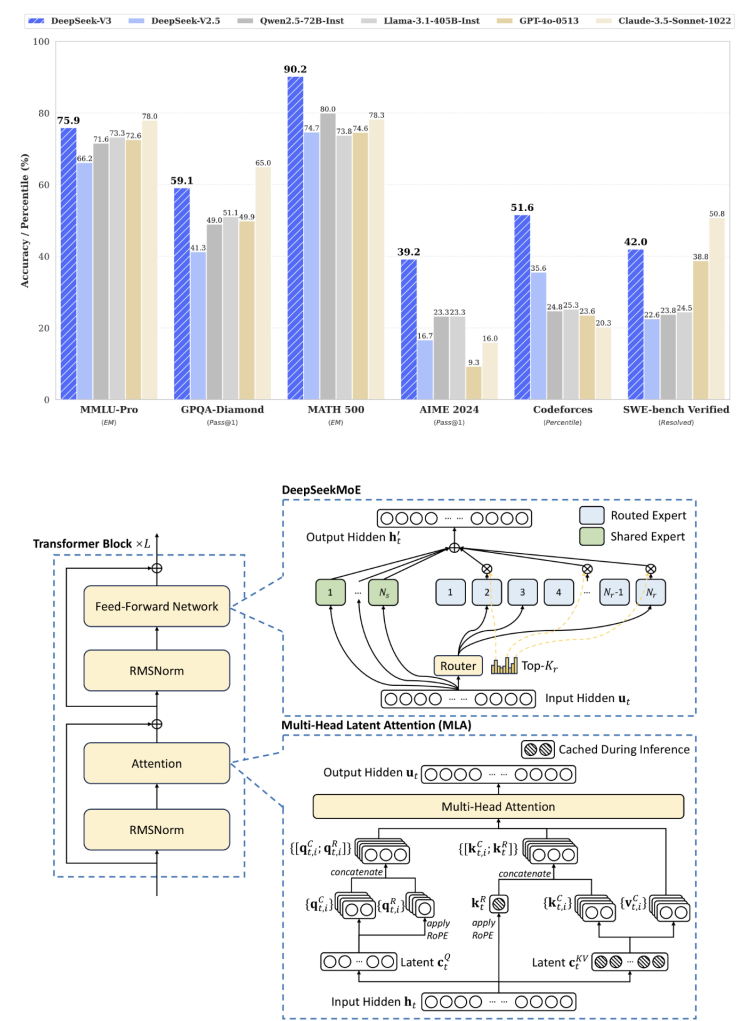
1. 架构创新：
 - 在DeepSeek-V2高效架构基础上，首创无辅助损失的负载均衡策略，显著减少了为实现负载均衡而带来的性能损失
 - 研究并验证了多Token预测(MTP)目标的有效性，该目标不仅提升模型性能，还可用于推理阶段的预测解码加速
 - 保持了DeepSeek-V2中验证过的**多头潜在注意力(MLA)**和**DeepSeekMoE架构**的优势
2. 高效预训练：
 - 开创性地设计了**FP8混合精度训练框架**，首次在超大规模模型上验证其可行性和有效性
 - 通过算法、框架和硬件的协同优化设计，成功克服了跨节点MoE训练中的通信瓶颈，实现了计算与通信的近乎完全重叠
 - 显著提升训练效率并降低成本，使模型规模扩展不会带来额外开销

- 仅使用2.664M H800 GPU小时就完成了14.8T tokens的预训练，后续训练阶段仅需0.1M GPU小时

3. 后训练知识蒸馏：

- 创新性地从DeepSeek-R1系列长链思维(CoT)模型中提取推理能力
- 设计了独特的管道将R1模型的验证和反思模式优雅地整合到DeepSeek-V3中
- 在显著提升推理性能的同时，成功维持了DeepSeek-V3的输出风格和长度控制
- 这种知识蒸馏方法为提升大语言模型的推理能力提供了新思路

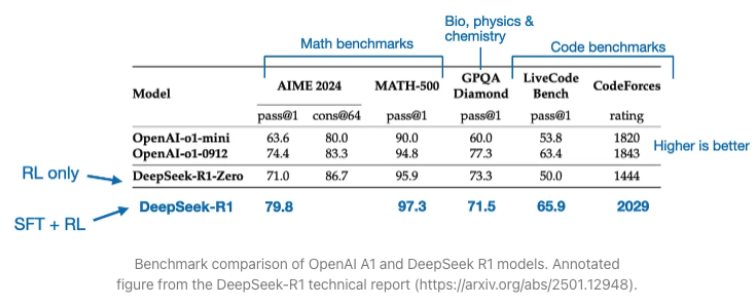
<https://arxiv.org/html/2412.19437v1>



1. DeepSeek-V3的架构设计体现了继承与创新的平衡，继承了DeepSeek-V2验证有效的MLA和DeepSeekMoE核心架构，同时通过引入MTP训练目标等创新设计提升模型性能，保持了基础架构的稳定性。

			<p>2. 架构设计实现了多维度优化，通过MLA机制实现高效推理，利用DeepSeekMoE架构降低训练成本，引入MTP提升模型性能，在性能、效率和经济性三个维度达成平衡。</p> <p>3. 整体设计展现了大模型架构的渐进优化思路，在保持核心优势的同时针对性解决特定问题，体现了模型迭代升级的合理路径。</p>
DeepSeek-R1-Zero			<p>从Deepseek-V3训练而来，纯强化学习（Pure RL）</p> <p>强化学习+两种奖励进行训练（称为“冷启动”训练，其不包括监督微调SFT）</p> <p>准确率奖励：利用 LeetCode 编译器来判定编程答案是否正确，以及用一套确定性系统来判定数学问题回答是否准确。</p> <p>格式奖励：用一个 LLM 评估器来判断回答的格式是否符合要求，比如推理步骤是否出现在 <code><think></code> 标签中等。</p> <p>Aha 模型开始在回答里自动生成推理线索，哪怕从来没有专门教它这样做 “顿悟”。</p> <div style="border: 1px solid black; padding: 5px; margin: 10px 0;"> <p>Question: If $a > 1$, then the sum of the real solutions of $\sqrt{a - \sqrt{a + x}} = x$ is equal to</p> <p>Response: <think></p> <p>To solve the equation $\sqrt{a - \sqrt{a + x}} = x$, let's start by squaring both ...</p> $(\sqrt{a - \sqrt{a + x}})^2 = x^2 \implies a - \sqrt{a + x} = x^2.$ <p>Rearrange to isolate the inner square root term:</p> $(a - x^2)^2 = a + x \implies a^2 - 2ax^2 + (x^2)^2 = a + x \implies x^4 - 2ax^2 - x + (a^2 - a) = 0$ <p>...</p> <p style="border: 1px solid red; padding: 2px; display: inline-block;">Wait, wait. Wait. That's an aha moment I can flag here.</p> <p>Let's reevaluate this step-by-step to identify if the correct sum can be ...</p> <p>We started with the equation:</p> $\sqrt{a - \sqrt{a + x}} = x$ <p>First, let's square both sides:</p> $a - \sqrt{a + x} = x^2 \implies \sqrt{a + x} = a - x^2$ <p>Next, I could square both sides again, treating the equation: ...</p> <p>...</p> </div> <p>存在的问题：过程混乱结果导向/中英文混杂</p>
DeepSeek-R1-0	2025.01.20	671B及多个蒸馏版本	<p>DeepSeek-R1 SFT + RL</p> <p>基于DeepSeek-R1-Zero进一步SFT和RL训练</p>

- 1. 基于R1-Zero生成冷启动的SFT数据（因为R1-Zero没有经历SFT）
- 2. 指令微调
- 3. 强化学习RL（准确性激励+格式激励+“语言一致性（应对中英文混杂的情况）”）
- 4. SFT数据再生成
 - a. 新第三步的模型构建带有CoT思维链的SFT数据，60万条
 - b. 使用Deepseek V3生成20万条通用知识的SFT数据
- 5. 强化学习RL（80万条数据），数学/编程问题使用规则问题判别（确定性答案），其他问题使用人类偏好打分



蒸馏版本

（通常小模型去学习大模型输出的logits，并结合原始目标数
据来训练）

这里是使用大模型生成的指令数据（SFT 数据）来直接微调小
模型（Qwen1.5B-32B, Llama 8B-70B）

使用的数据同样是 DeepSeek-R1 与 DeepSeek-V3 产生 SFT
数据

和 DeepSeek-R1 相比，蒸馏模型弱一些，但对比 R1-Zero 表
现不错，而且在规模上要小得多。

部分模型也优于 o1 mini （是否也是蒸馏？）

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
DeepSeek-R1-Zero	71.0		95.9	73.3	50.0	1444
DeepSeek-R1	79.8		97.3	71.5	65.9	2029

同时R1还做一些其他尝试，如果直接在小模型上做RL是否有好的效果？

Qwen-32B 也和 R1-Zero 的纯 RL 思路来训练，对于较小的模型，**纯 RL** 效果不如“蒸馏”。纯RL 去训练数十亿参数级的模型，推理能力并不容易实现。

Model	AIME 2024		MATH-500	GPQA Diamond	LiveCodeBench
	pass@1	cons@64	pass@1	pass@1	pass@1
pure RL → 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
pure RL → DeepSeek-R1-Zero	71.0		95.9	73.3	50.0
DeepSeek-R1	79.8		97.3	71.5	65.9

Janus-Pro

2025.01.28

1B/7B

多模态融合模型，支持理解和生成

https://github.com/deepseek-ai/Janus/blob/main/janus_pro_tech_report.pdf