

Artificial Intelligence Newsletter

人工智能 前沿快报

2025 年第 01 期
总第 19 期



学会官方公众号

广东省图象图形学会

GDSIG

学术交流、技术咨询、科学普及

目录

学术动态

一. NOISE_STEP: TRAINING IN 1.58B WITH NO GRADIENT MEMORY.....	1
二. Scaling of Search and Learning: A Roadmap to Reproduce o1 from Reinforcement Learning Perspective...2	2
三. Memory Layers at Scale.....	4
四. LongDocURL: a Comprehensive Multimodal Long Document Benchmark Integrating Understanding, Reasoning, and Locating.....	6
五. 1.58-bit FLUX.....	8
六. Next Token Prediction Towards Multimodal Intelligence: A Comprehensive Survey.....	10
七. Titans: Learning to Memorize at Test Time.....	12
八. Metadata Conditioning Accelerates Language Model Pre-training.....	14
九. 2.5 Years in Class: A Multimodal Textbook for Vision-Language Pretraining.....	16
十. VITA-1.5: Towards GPT-4o Level Real-Time Vision and Speech Interaction.....	18
十一. A Survey on Large Language Models with some Insights on their Capabilities and Limitations.....	20
十二. Test-time Computing: from System-1 Thinking to System-2 Thinking.....	22
十三. From Generation to Judgment: Opportunities and Challenges of LLM-as-a-judge.....	24
十四. Large language models for artificial general intelligence (AGI): A survey of foundational principles and approaches.....	26
十五. Dolphin: Closed-loop Open-ended Auto-research through Thinking, Practice, and Feedback.....	28
十六. Accurate predictions on small data with a tabular foundation model.....	30
十七. On Computational Limits and Provably Efficient Criteria of Visual Autoregressive Models: A Fine-Grained Complexity Analysis.....	32
十八. rStar-Math: Small LLMs Can Master Math Reasoning with Self-Evolved Deep Thinking.....	33
十九. The Future of AI: Exploring the Potential of Large Concept Models.....	35
二十. Toward expert-level medical question answering with large language models.....	37
二十一. LlamaV-o1: Rethinking Step-by-step Visual Reasoning in LLMs.....	39
二十二. The Lessons of Developing Process Reward Models in Mathematical Reasoning.....	42
二十三. Transformer2: Self-adaptive LLMs.....	44
二十四. OmniThink: Expanding Knowledge Boundaries in Machine Writing through Thinking.....	46
二十五. A generative model for inorganic materials design.....	48
二十六. LLM360 K2: Building a 65B 360-Open-Source Large Language Model from Scratch.....	50
二十七. Mobile-Agent-E: Self-Evolving Mobile Assistant for Complex Tasks.....	52

目录

二十八. DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning.....	54
--	----

技术动态

一、上海人工智能实验室进一步开源 InternVL2.5 后训练算法及训练数据.....	57
二、Sam Altman 新年博文：Reflections，预见 AGI 时代已经到来，相信超级智能（superintelligence）工具可以大大加速科学发现和创新.....	58
三、《MIT Technology Review》：What’s next for AI in 2025.....	60
四、Sky-T1: Train your own O1 preview model within \$450.....	61
五、MiniMax-01: Scaling Foundation Models with Lightning Attention.....	59
六、上海 AI Lab 发布 Intern3-8B-Instruct 模型.....	63
七、OpenAI 发布首个 L3 级智能体.....	64
八、China’s cheap, open AI model DeepSeek thrills scientists.....	65

资源分享

一、Foundations of Large Language Models.....	68
---	----

INTRODUCTION

导读

在人工智能技术飞速发展的时代背景下，我们将不定期梳理人工智能领域的部分前沿学术与技术动态，并以简报的形式分享给读者，以帮助关注此领域的人士了解最新的学术前沿发展与产业技术动态，促进学术交流和技術繁荣。本期摘选了近期全球人工智能领域的 28 篇最新学术动态、8 篇技术动态、以及 1 个资源推荐给广大读者。

ACADEMIC TRENDS

学术动态

● 论文一

NOISE_STEP: TRAINING IN 1.58B WITH NO GRADIENT MEMORY

日期

2024-12-08

NOISE_STEP: TRAINING IN 1.58B WITH NO GRADIENT MEMORY

PREPRINT

Will Brickner
wgbrickner@gmail.com

December 23, 2024

内容摘要

本文提出了一种新的算法，旨在通过直接在三进制精度（1.58 比特）下进行训练，解决当前大型机器学习模型训练的计算和能源负担问题。该算法引入了雅可比向量积（Jacobian Vector Product, JVP）这种不依赖

反向传播的梯度估计方法，能够与模型推理并行进行训练，极大地减少了内存和能源消耗。此外，通过噪声步（noise_step）方法，模型可以在分布式训练中更高效地运行，并且使用非常紧凑的步骤表示，从而减少了模型传输的存储需求。该方法的一个显著特点是通过梯度估计和三进制表示来降低训练和通信开销。

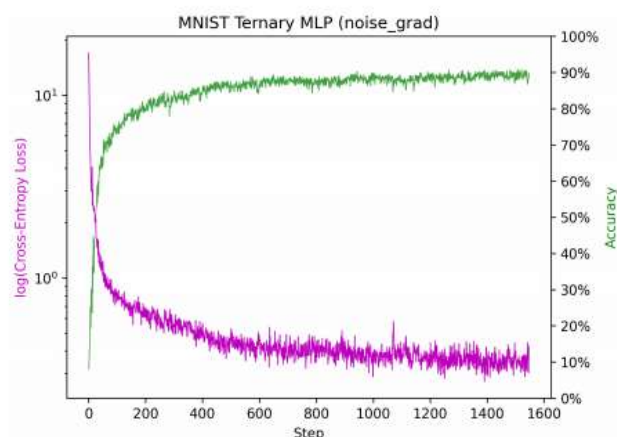


Figure 1: noise_step loss and accuracy curves
 $\{ \text{samples}=128, \text{density}=6 \times 10^{-5} \rightarrow 3.7 \times 10^{-6} \}$

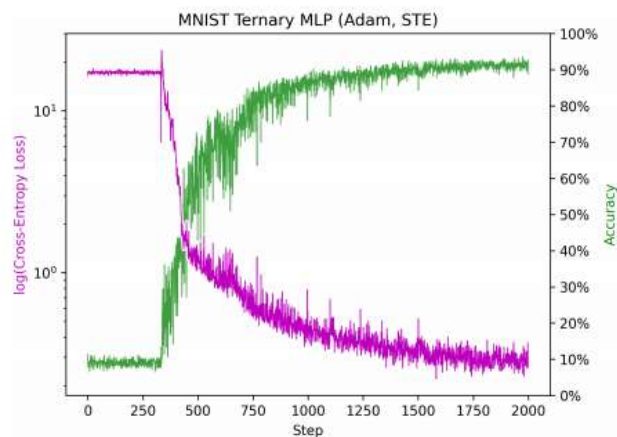


Figure 2: Adam optimizer loss and accuracy curves, default parameters.

主要亮点

- 提出了直接在三进制精度下进行训练的算法，大幅减少训练过程中所需的内存和能源消耗。
- 无需传统的反向传播或动量，能够与模型推理并行执行，降低计算资源需求。
- 提供了一种高效的分布式训练方法，显著减少了通信开销，适用于大规模模型的训练。
- 通过简化模型更新和步骤表示，大幅度减小了模型的存储需求。

相关链接

Paper: https://github.com/wbrickner/noise_step/blob/main/latex/noise_step.pdf

● 论文二

Scaling of Search and Learning: A Roadmap to Reproduce o1 from Reinforcement Learning Perspective

日期

2024-12-18

Scaling of Search and Learning: A Roadmap to Reproduce o1 from Reinforcement Learning Perspective

Zhiyuan Zeng^{1*} Qinyuan Cheng^{1*} Zhangyue Yin^{1*} Bo Wang^{1*}
Shimin Li¹ Yunhua Zhou² Qipeng Guo² Xuanjing Huang¹ Xipeng Qiu^{1†}

¹Fudan University

²Shanghai AI Laboratory

内容摘要

本文提出了一条基于强化学习视角的道路图，以复现 OpenAI o1 模型的强大推理能力为目标，深入分析了四个关键组件：策略初始化（Policy Initialization）、奖励设计（Reward Design）、搜索（Search）和学习（Learning）。这些组件共同支撑了 o1 模型的推理能力，使其能够执行复杂的任务，如问题分解、自我校正以及生成高质量的解决方案。论文通过分析现有开源项目和理论框架，为构建具有强大推理能力的 o-1 类大语言模型（LLM）提供了系统性指导。

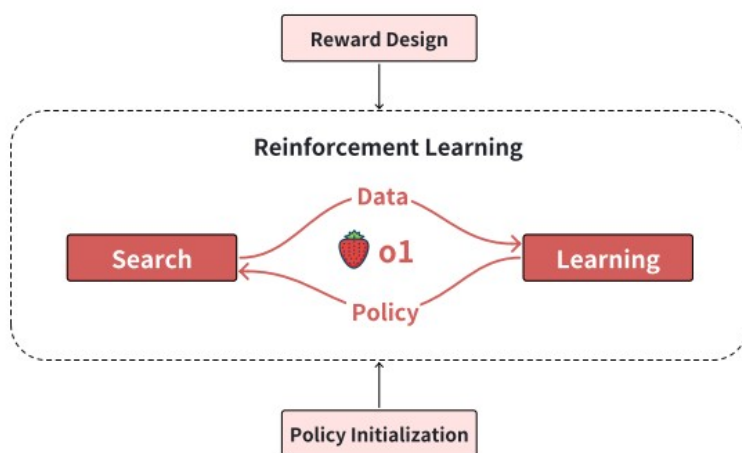


Figure 1: The overview of this roadmap including policy initialization, reward design, search and learning.

主要亮点

- 提出了一条以强化学习为核心的道路图，涵盖了策略初始化、奖励设计、搜索和学习四个关键方面，为复现 o1 模型的推理能力奠定了理论基础。

- 深入探讨了策略初始化过程，强调通过预训练和指令微调获取类人推理行为，并结合示例任务明确了模型探索解空间的必要性。
- 提出了细粒度的奖励设计方法，包括结果奖励和过程奖励，提供了对稀疏奖励信号进行重构的指导性建议，如奖励模型的使用和潜在奖励设计。
- 分析了搜索在训练和推理阶段的关键作用，提出基于内部和外部指导信号的搜索策略，如树搜索（Tree Search）和序列修正（Sequential Revisions），为复杂任务提供解决方案。
- 强调了学习模块中强化学习与行为克隆的结合，通过数据交互和策略改进实现更优性能，并参考 AlphaGo 的成功案例展示了搜索与学习的协同作用。
- 对现有开源项目进行了总结，指出它们如何成为道路图中的组成部分或变体，并探讨了未来发展趋势与挑战，为构建类似 o1 的模型提供了实践性指导。

相关链接

Paper: <https://arxiv.org/pdf/2412.14135>

● 论文三

Memory Layers at Scale

日期

2024-12-23

Memory Layers at Scale

Vincent-Pierre Berges*, Barlas Oğuz*, Daniel Haziza, Wen-tau Yih, Luke Zettlemoyer, Gargi Ghosh

Meta FAIR

*Main authors

内容摘要

本文提出了一种新的内存层方法，通过在现有模型中集成一个可训练的键值查找机制，增强了模型的存储与信息检索能力，且在计算上保持高效。与计算密集型的全连接层不同，稀疏激活的内存层为模型提供了专门的存储容量，并能够高效地执行信息检索。本文通过扩展内存层并在实际规模下进行测试，证明了其在各种下游任务中的有效性，尤其是在事实性任务上，显著提高了语言模型的表现。实验结果表明，增强内存层的模型相比于传统的全连接网络，能够在较少计算量的条件下，提升多达两倍的性能。本文还提供了一个完全并行化的内存层实现，展示了其在最大 1280 亿内存参数规模下的扩展性。

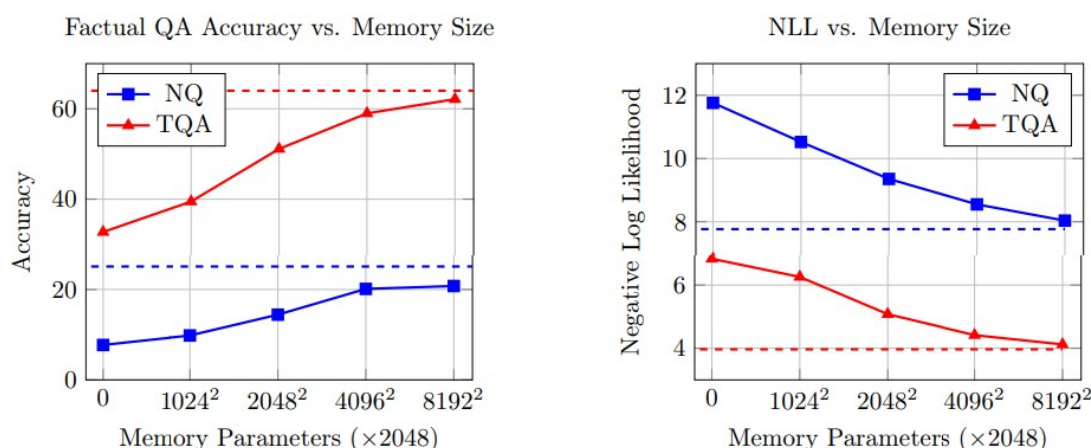


Figure 1 Scaling the size of the memory for a 1.3 billion parameter base model (zero memory parameters corresponds to a dense model), trained to 1 trillion tokens. On the left, factual QA accuracy (exact match on NaturalQuestions and F1 score on TriviaQA), on the right task NLL (lower is better). Dashed lines show the performance of a 7B model trained on 2 trillion tokens with 10x more FLOPs.

主要亮点

- 提出了增强内存层的概念，并展示了其在大语言模型中提高任务性能的潜力，尤其在事实性问题回答中表现显著。
- 内存层通过替代 Transformer 中的前馈层（FFN），在不增加计算量的情况下，显著提高了模型的性能和存储效率。

- 提供了一个可扩展的内存层实现，支持最大 1280 亿内存参数，并能高效训练到 1 万亿 tokens，表现出优秀的扩展性。
- 内存层相比于密集模型和专家混合模型（MOE），在事实性任务上具有明显优势，尤其在问答任务中表现突出。
- 在硬件优化方面，提出了高效的内存查找机制和并行化实现，提升了内存操作的效率，并显著加速了训练过程。

相关链接

Paper: <https://arxiv.org/pdf/2412.09764>

Code: <https://github.com/facebookresearch/memory>

Blog: <https://ai.meta.com/blog/meta-fair-updates-agents-robustness-safety-architecture>

● 论文四

LongDocURL: a Comprehensive Multimodal Long Document Benchmark Integrating Understanding, Reasoning, and Locating

日期

2024-12-24

LongDocURL: a Comprehensive Multimodal Long Document Benchmark Integrating *Understanding, Reasoning, and Locating*

Chao Deng^{§*†}, Jiale Yuan^{◇*}, Pi Bu[◇], Peijie Wang[§], Zhong-Zhi Li[§], Jian Xu[§],
Xiao-Hui Li[§], Yuan Gao[◇], Jun Song^{◇‡}, Bo Zheng[◇], Cheng-Lin Liu^{§‡}

MAIS, Institute of Automation of Chinese Academy of Sciences[§]

School of Artificial Intelligence, University of Chinese Academy of Sciences[§], Alibaba Group[◇]

{dengchao2023, wangpeijie2023, lizhongzhi2022, jian.xu}@ia.ac.cn[§],

{yuanjiale.yjl, bupi.wj, jsong.sj, bozheng}@alibaba-inc.com[◇],

{xiaohui.li, liucl}@nlpr.ia.ac.cn[§]

内容摘要

本文提出了 LongDocURL，一个全面的多模态长文档基准，旨在解决现有文档理解基准无法处理复杂文档元素和长篇内容的问题。LongDocURL 结合了三个主要任务：文档理解、数值推理和跨元素定位。该基准包含 20 个子任务，涵盖不同的文档元素和任务类型，目标是提供对长文档和复杂任务的全面评估。为了构建该基准，团队设计了一个半自动化的数据构建管道，从 33,000 多页的文档中收集了 2,325 个高质量问答对。通过在多个开源和闭源模型上进行评估，本文揭示了当前模型在处理长文档时的性能差距，并提供了 LongDocURL 测试基准作为进一步研究和改进文档理解模型的参考标准。

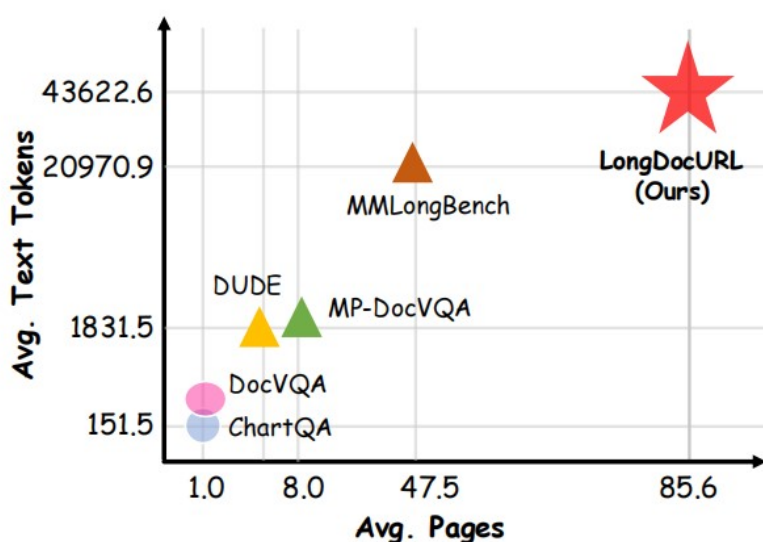


Figure 1: Comparison with other datasets in **average pages and text tokens per document**.

主要亮点

- 提出了 LongDocURL 基准，涵盖了文档理解、数值推理和跨元素定位三个主要任务，包含 20 个子任务，支持更细粒度的评估。
- 设计并实现了一个高效的半自动化数据构建管道，生成了覆盖 33,000 多页文档的 2,325 个高质量问答对，显著超越现有基准。

- 对开源和闭源大模型进行了全面的评估，揭示了当前模型在多层文档理解中的性能差距，展示了该领域亟待改进的地方。
- 通过数据集的统计分析，提供了关于不同文档源、任务类型和文档元素的细粒度评估，指明了未来文档模型的优化方向。
- 强调了表格结构和布局信息在多模态长文档任务中的重要性，尤其是对开放源代码模型的影响。

相关链接

Paper: <https://arxiv.org/pdf/2412.18424>

Github: <https://github.com/dengc2023/LongDocURL>

● 论文五

1.58-bit FLUX

日期

2024-12-24

1.58-bit FLUX

Chenglin Yang¹, Celong Liu¹, Xueqing Deng¹, Dongwon Kim²,
Xing Mei¹, Xiaohui Shen¹, Liang-Chieh Chen¹

¹ByteDance ²POSTECH

<https://chenglin-yang.github.io/1.58bit.flux.github.io/>

内容摘要

本文介绍了 1.58-bit FLUX，这是第一个成功地将 FLUX 文本到图像生成模型进行 1.58 位量化的方法。该方法通过将模型的视觉变换器（Vision Transformer）权重限制为三种值（-1, 0, +1），在不依赖于图像数据的情况下实现了极其低位的量化，从而显著降低了模型的存储需求。通过引入专门优化的 1.58 位计算内核，

1.58-bit FLUX 在减少模型存储和推理内存使用的同时，还提高了推理速度。大量的实验评估表明，1.58-bit FLUX 在多个基准测试中表现出与全精度 FLUX 相当的生成质量，并显著提升了计算效率。

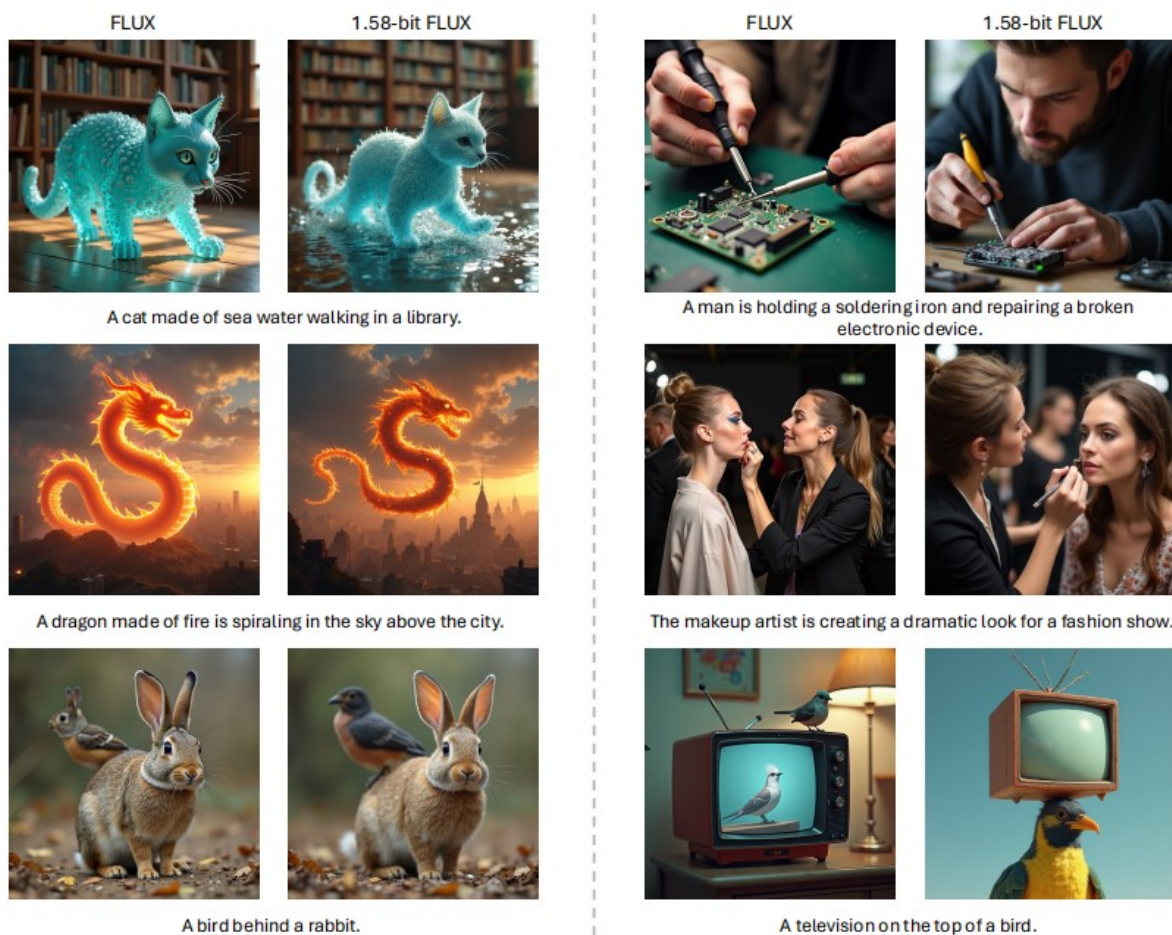


Figure 1. **Visual comparisons between FLUX and 1.58-bit FLUX.** 1.58-bit FLUX demonstrates comparable generation quality to FLUX while employing 1.58-bit quantization, where 99.5% of the 11.9B parameters in the vision transformer are constrained to the values +1, -1, or 0. For consistency, all images in each comparison are generated using the same latent noise input. 1.58-bit FLUX utilizes a custom 1.58-bit kernel. Additional visual comparisons are provided in Fig. 3 and Fig. 4.

主要亮点

- 提出了 1.58-bit FLUX，首次通过 1.58 位量化 FLUX 视觉变换器的 99.5% 参数，显著降低了存储需求。
- 开发了一种针对 1.58 位计算优化的高效内核，带来了 7.7 倍的模型存储压缩和 5.1 倍的推理内存减少，同时提高了推理延迟。

- 实验结果表明, 1.58-bit FLUX 在 T2I (文本到图像) 生成的多个基准测试中, 与全精度 FLUX 模型相比, 生成质量没有显著下降。
- 在资源受限的设备 (如移动设备) 上, 1.58-bit FLUX 展现了更好的部署效率, 提供了一个高效的解决方案。
- 该方法不依赖图像数据, 只使用自监督机制进行量化, 扩展了量化方法在生成模型中的应用范围。

相关链接

Paper: <https://arxiv.org/pdf/2412.18653>

Project: <https://chenglin-yang.github.io/1.58bit.flux.github.io/>

Code: <https://github.com/Chenglin-Yang/1.58bit.flux>

● 论文六

Next Token Prediction Towards Multimodal Intelligence: A Comprehensive Survey

日期

2024-12-16

Next Token Prediction Towards Multimodal Intelligence: A Comprehensive Survey

LIANG CHEN^{1†} ZEKUN WANG^{2*} SHUHUAI REN^{1*} LEI LI^{3*} HAOZHE ZHAO^{1*} YUNSHUI LI^{4*} ZEFAN CAI¹ HONGCHENG GUO² LEI ZHANG⁴ YIZHE XIONG⁵ YICHI ZHANG¹ RUOYU WU¹ QINGXIU DONG¹ GE ZHANG⁶ JIAN YANG⁸ LINGWEI MENG⁷ SHUJIE HU⁷ YULONG CHEN⁹ JUNYANG LIN⁸ SHUAI BAI⁸ ANDREAS VLACHOS⁹ XU TAN¹⁰ MINJIA ZHANG¹¹ WEN XIAO¹⁰ AARON YEE^{12,13} TIANYU LIU⁸ BAOBAO CHANG¹

内容摘要

本论文综述了面向多种模态的下一个标记预测 (Next Token Prediction, NTP) 方法的研究进展。NTP 最初用于自然语言处理中的语言建模任务, 但随着大规模语言模型 (LLMs) 的发展, NTP 被扩展到处处理多种模

态的数据，如图像、视频、音频等。通过将不同模态的信息转化为标记并在上下文中预测下一个标记，NTP 能够同时处理理解和生成任务。论文提出了一种综合的分类方法，涵盖了多模态标记化、MMNTP 模型架构、统一任务表示、数据集与评估以及面临的挑战。该分类方法旨在为研究人员提供一个框架，帮助他们更加好地探索多模态智能。

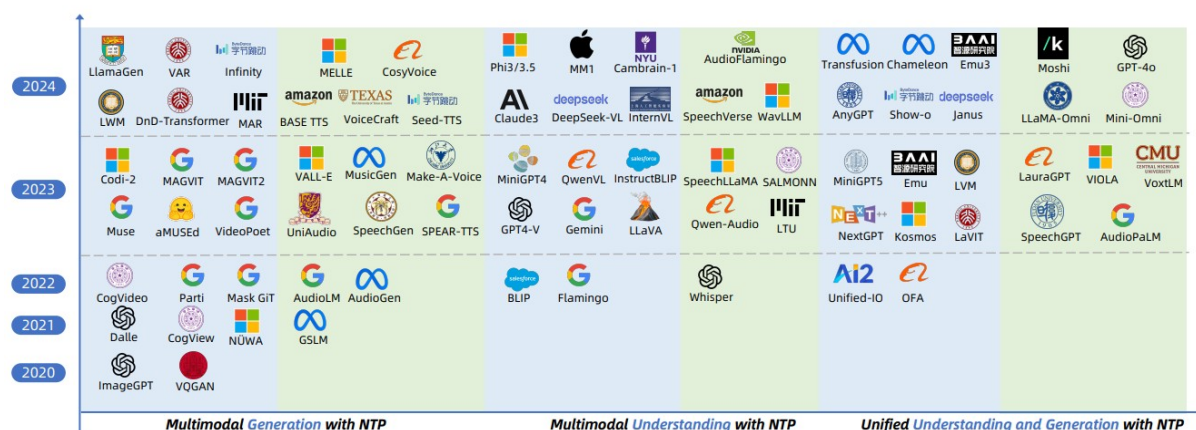


Fig. 1. Historical development of LMMs utilizing Next-Token Prediction. Models featuring vision and more modalities are set in blue background while models that support audio modality are set in green backgrounds.

主要亮点

- 提出了多模态下的 NTP 任务，统一了不同模态下的理解与生成任务，成为多模态学习的核心方法。
- 介绍了两种主要的标记化方法——离散标记化与连续标记化，以及它们在图像、视频、音频等模态中的应用。
- 讨论了多模态模型的架构，包括如何通过 Transformer 等架构处理不同模态的数据。
- 论文总结了多个开源数据集和评估标准，并指出了 NTP 模型在多模态任务中的优势与挑战。

相关链接

Paper: <https://arxiv.org/pdf/2412.18619>

Github: <https://github.com/LMM101/Awesome-Multimodal-Next-Token-Prediction>

● 论文七

Titans: Learning to Memorize at Test Time

日期

2024-12-31

Titans: Learning to Memorize at Test Time

Ali Behrouz[†], Peilin Zhong[†], and Vahab Mirrokni[†]

[†]Google Research

{alibehrouz, peilinz, mirrokni}@google.com

内容摘要

论文提出了一种名为 Titans 的新型深度学习架构，通过设计神经长期记忆模块（Neural Long-Term Memory Module, LMM），解决了 Transformer 和现代线性 RNN 在长序列处理、泛化能力与记忆瓶颈上的局限性。Titans 架构结合短期记忆（Attention）和长期记忆（LMM），能够在测试时动态学习、记忆和遗忘信息，有效提升模型在语言建模、常识推理、时间序列预测、DNA 建模等任务中的表现。此外，Titans 在超长上下文窗口（>2M tokens）任务中表现优于现有的主流模型（如 Transformer 和现代线性 RNN），展现了强大的扩展性和准确性。

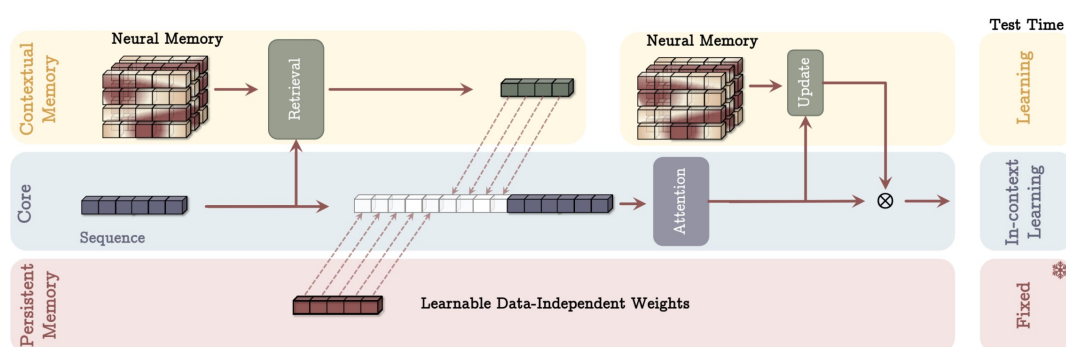


Figure 2: **Memory as a Context (MAC) Architecture.** This architecture includes three branches of (1) core, (2) contextual (long-term) memory, and (3) persistent memory. The core branch concatenates the *corresponding* long-term and persistent memories with the input sequence. Next, attention performs on the sequence and decides what part of the information should store in the long-term memory. At the test time, parameters corresponds to contextual memory are still learning, parameters corresponds to the core branch are responsible for in-context learning, and parameters of persistent memory are responsible to store the knowledge about tasks and so are fixed.

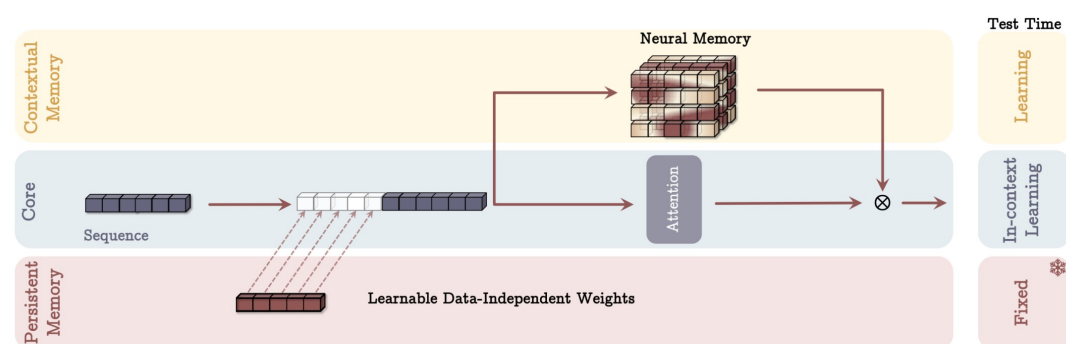


Figure 4: **Memory as a Gate (MAG) Architecture.** This architecture, similarly, has the three branches of (1) core, (2) contextual memory, and (3) persistent memory. It, however, incorporates only persistent memory into the context and combine memory with the core branch using a gating mechanism. At test time, the behavior is the same as Figure 2.

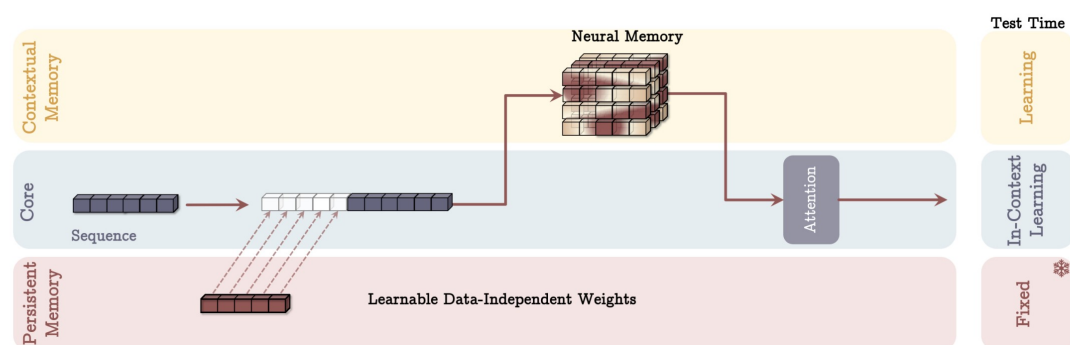


Figure 5: **Memory as a Layer (MAL) Architecture.** In this architecture, the memory layer is responsible to compress the past and current context before the attention module.

主要亮点

- 提出神经长期记忆模块 (LMM)：设计了一种神经长期记忆模块，能够通过动态记忆和遗忘机制在测试时学习和更新记忆，同时基于“惊喜”度量优化记忆存储，显著提升泛化能力。

- 多变体架构设计：提出了三种 Titans 架构变体（Memory as Context, Memory as Gating, Memory as a Layer），分别将长期记忆模块作为上下文、门控机制或独立层，满足不同任务需求，展示了性能与效率的权衡。
- 长序列建模能力：Titans 架构在超长上下文窗口（>2M tokens）任务中表现出色，尤其在“大海捞针”任务（Needle-in-Haystack）和复杂推理任务中优于 GPT-4 和其他 SOTA 模型。
- 深度记忆模块优势：通过实验验证深度记忆模块（Deep Memory）的有效性，发现更深的记忆模块能显著提升模型在长序列任务中的性能，同时展示了深度与效率的平衡点。
- 广泛应用与高效性：Titans 在语言建模、常识推理、时间序列预测和 DNA 建模等多领域任务中均优于现有模型，同时支持高效的训练与推理，展示了出色的计算效率与通用性。

相关链接

Paper: <https://arxiv.org/pdf/2501.00663>

● 论文八

Metadata Conditioning Accelerates Language Model Pre-training

日期

2025-1-3

Metadata Conditioning Accelerates Language Model Pre-training

Tianyu Gao Alexander Wettig Luxi He Yihe Dong Sadhika Malladi Danqi Chen
Princeton Language and Intelligence, Princeton University
tianyug@princeton.edu

内容摘要

本文提出了一种新的方法——元数据调节（Metadata Conditioning then Cooldown，简称 MeCo），用于加速语言模型的预训练。传统的预训练数据来源多样，涵盖不同的领域和质量，处理这些异质数据具有挑战性。MeCo 通过在训练过程中将元数据（如网页 URL）与文本一起输入，帮助模型在训练过程中更好地理解 and 利用这些数据的背景信息。该方法显著提高了预训练效率，减少了所需的训练数据量，且通过“冷却”阶段确保模型能够在推理时无须元数据。MeCo 不仅提升了预训练的效率，还使得模型能够在推理过程中通过条件化推理，利用合适的元数据来引导模型生成更为符合预期的内容。实验结果表明，使用 MeCo 的模型在多个任务中表现出色，并且能够减少有害生成内容。

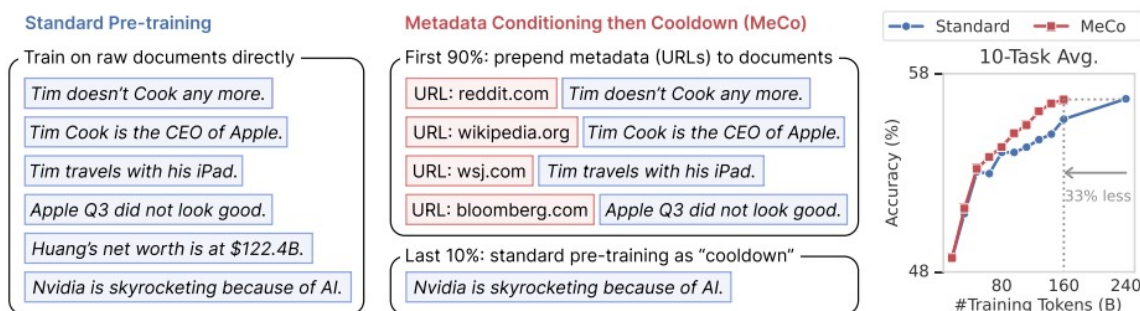


Figure 1: A comparison between data used by standard pre-training and MeCo. The figure on the right demonstrates 5-shot downstream task performance averaged across 10 tasks (1.6B models; details about the experiments can be found in §3).

主要亮点

- 提出了 MeCo 方法，通过在预训练阶段加入元数据调节，显著提高语言模型的训练效率，减少 33% 的数据使用量。
- MeCo 能够通过“冷却”阶段确保模型在推理时无需元数据，提升了模型的通用性和稳定性。
- 通过条件化推理，MeCo 允许通过预先加入的真实或虚构的元数据来引导模型生成符合要求的内容，改善了模型的可控性和任务表现。

- 在多个模型规模（从 600M 到 8B 参数）和数据集（如 C4、RefinedWeb、DCLM）上，MeCo 均表现出了较好的性能提升，且其方法简单，未增加额外计算开销。
- MeCo 还通过减少有害生成内容的可能性，提升了模型在处理敏感话题时的安全性，尤其是在使用如 Wikipedia.org 这样的元数据时，能有效减少毒性生成。

相关链接

Paper: <https://arxiv.org/pdf/2501.01956>

Github: <https://github.com/princeton-pli/MeCo>

● 论文九

2.5 Years in Class: A Multimodal Textbook for Vision-Language Pretraining

日期

2025-1-3



2.5 Years in Class: A Multimodal Textbook for Vision-Language Pretraining

Wenqi Zhang^{1*} Hang Zhang² Xin Li^{2,†} Jiashuo Sun²
Yongliang Shen¹ Weiming Lu^{1,†} Deli Zhao² Yueting Zhuang¹ Lidong Bing²

¹College of Computer Science and Technology, Zhejiang University

²DAMO Academy, Alibaba Group

zhangwenqi@zju.edu.cn

Project: <https://multimodal-interleaved-textbook.github.io/>

内容摘要

本文提出了一种高质量的多模态教材，用于视觉-语言模型（VLM）的预训练。该教材基于超过 2.5 年的教学视频（共 22,000 课时），涵盖数学、物理等六个基础学科，旨在提供丰富且系统的知识，使模型能够以更自然的方式理解图文结合的内容。与传统的网页抓取数据集相比，本文通过逐步提取视频中的视觉关键帧、自

动语音识别 (ASR) 文本和光学字符识别 (OCR) 文本, 将这些信息按时间顺序组织成图文交错的格式。实验结果显示, 在多个知识密集型和推理任务 (如 ScienceQA 和 MathVista) 上, 基于该教材进行预训练的模型表现优异, 特别在任务解决中展现了更强的图文结合上下文理解能力。

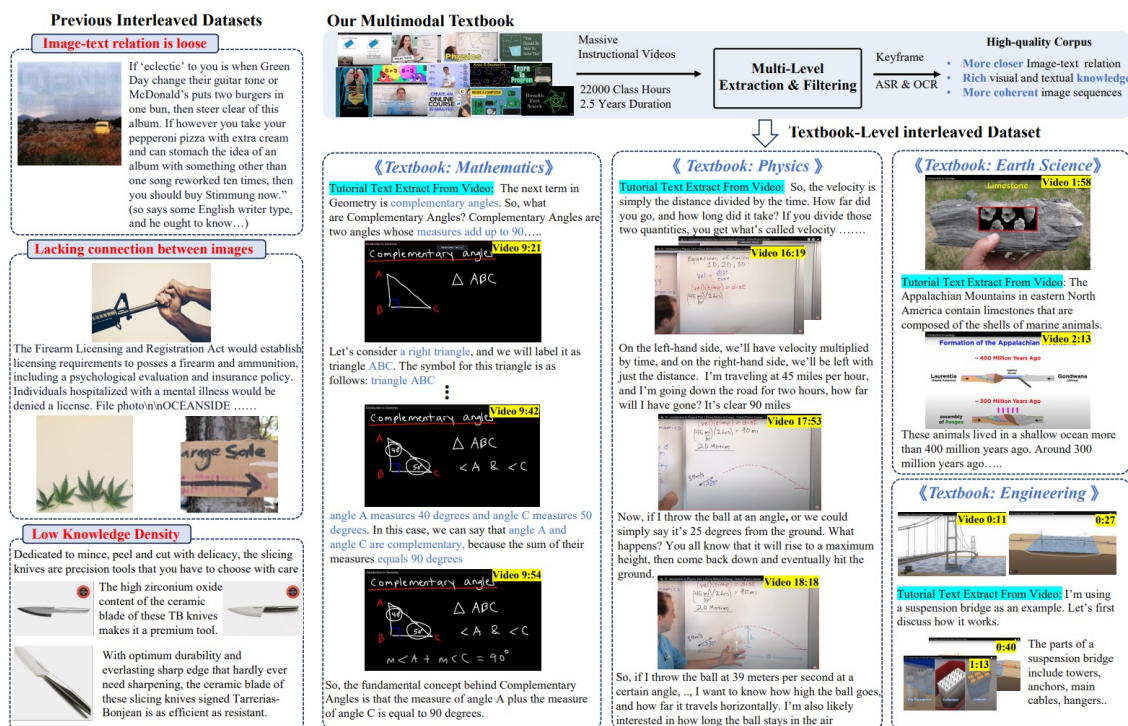


Figure 1. Previous interleaved datasets, e.g., MMC4 and OBELICS, suffer from limitations like weak text-image relations, low knowledge density, and incoherent image sequences. Our multimodal textbook, sourced from massive tutorial videos, employs coarse-to-fine knowledge extraction and multi-level filtering to create a high-quality, textbook-level dataset. It interleaves video keyframes with tutorial texts (extracted from ASR and OCR), enabling VLMs to acquire rich knowledge through tightly coupled text-image and more coherent logic.

主要亮点

- 提出了一个多模态教材, 基于超过 2.5 年的教学视频, 创建了一个高质量的图文交错数据集, 具有更高的知识密度和逻辑连贯性。
- 开发了一个多层次的知识提取和数据筛选管道, 通过视觉关键帧提取、ASR 和 OCR 文本提取, 确保教材数据的高质量。
- 在多个基准任务 (如 ScienceQA、MathVista) 上, 使用该教材进行预训练的 VLM 模型在知识和推理密集型任务上取得了显著提升。

- 该教材改善了 VLMs 在少样本学习中的图文结合上下文意识，使模型能够更有效地利用视觉和文本提示来解决任务。
- 相较于其他现有的数据集，本文的多模态教材提供了更一致的图文关联和更紧密的逻辑关系，使 VLM 在处理复杂任务时表现更好。

相关链接

Paper: <https://arxiv.org/pdf/2501.00958>

Project: <https://multimodal-interleaved-textbook.github.io/>

Github: https://github.com/DAMO-NLP-SG/multimodal_textbook

Dataset: https://huggingface.co/datasets/DAMO-NLP-SG/multimodal_textbook

● 论文十

VITA-1.5: Towards GPT-4o Level Real-Time Vision and Speech Interaction

日期

2025-1-3

VITA-1.5: Towards GPT-4o Level Real-Time Vision and Speech Interaction

Chaoyou Fu^{1,♣}, Haojia Lin³, Xiong Wang², Yi-Fan Zhang⁴, Yunhang Shen²
Xiaoyu Liu¹, Yangze Li², Zuwei Long², Heting Gao², Ke Li², Long Ma²,
Xiawu Zheng³, Rongrong Ji³, Xing Sun^{2,†}, Caifeng Shan¹, Ran He⁴

¹NJU, ²Tencent Youtu Lab, ³XMU, ⁴CASIA

♣ Project Leader † Corresponding Author

Demo Video: [Click YouTube Link](#)

Source Code: <https://github.com/VITA-MLLM/VITA>

内容摘要

本文提出了 VITA-1.5，一个集成了视觉、语言和语音能力的多模态大语言模型（MLLM）。VITA-1.5 采用了精心设计的三阶段训练方法，逐步将视觉和语音数据引入模型，从而缓解不同模态间的冲突，最终实现了流畅的视觉和语音交互。该方法通过训练视觉适配器和语音编码器，不仅保留了强大的视觉语言能力，还使模型能够在没有单独的语音识别（ASR）和语音合成（TTS）模块的情况下高效地实现语音对话。这一改进显著提高了多模态端到端响应速度，并在多个基准测试中展示了其在图像、视频和语音任务上的强大能力。

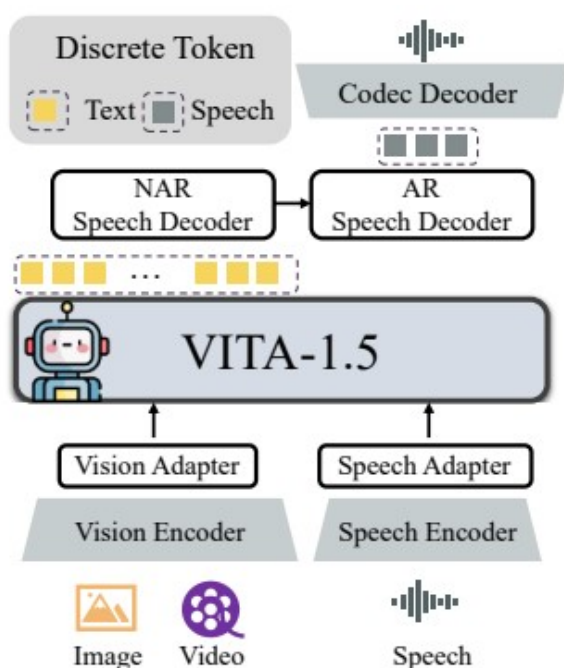


Figure 2: **Overall Architecture of VITA-1.5.** The input side consists of vision and audio encoders, along with their adapters connected to a LLM. The output side has an end-to-end speech generation module, rather than directly using an external TTS model as the initial VITA-1.0 version [16].

主要亮点

- 提出了 VITA-1.5 模型，通过三阶段训练策略，成功集成视觉、语言和语音模态，解决了多模态数据间的冲突，确保了强大的多模态理解能力。

- 模型能够实现无需外部 ASR 和 TTS 模块的语音到语音对话，显著提升了交互效率。
- 在多种基准测试中，VITA-1.5 展示了与现有先进开源和闭源模型相当的图像和语音理解能力，特别是在语音能力方面取得了显著进展。
- 引入了三阶段训练策略：首先是视觉语言训练，其次是语音输入微调，最后是语音输出微调，从而实现了高效的多模态互动。
- 实验结果表明，VITA-1.5 在图像理解、视频分析和语音识别等任务中表现出色，并且在与专有模型的比较中，取得了有竞争力的结果。

相关链接

Paper: <https://arxiv.org/pdf/2501.01957>

Github: <https://github.com/VITA-MLLM/VITA>

● 论文十一

A Survey on Large Language Models with some Insights on their Capabilities and Limitations

日期

2025-1-3

A Survey on Large Language Models with some Insights on their Capabilities and Limitations

Andrea Matarazzo
Expedia Group
Italy
a.matarazzo@gmail.com

Riccardo Torlone
Roma Tre University
Italy
riccardo.torlone@uniroma3.it

内容摘要

本文是一篇长达 174 页的大型语言模型（LLMs）综述，深入探讨分析了现有大模型的能力和局限性。基于 Transformer 架构的大语言模型，如 GPT、BERT 和 LLaMA，已经重新定义了自然语言处理的能力，能够在文本生成、问答、翻译、总结等任务中表现出色。更重要的是，LLMs 还表现出了超越其核心功能的能力，例如常识推理、代码生成和数学运算等。本文回顾了 LLMs 的基础组件、扩展机制和架构策略，分析了数据和计算增长对 LLMs 性能的影响，同时也探讨了模型扩展所面临的权衡。论文还探讨了 LLMs 在多个领域（如医疗、金融、教育和法律）的应用，揭示了这些模型在特定领域解决挑战的潜力。

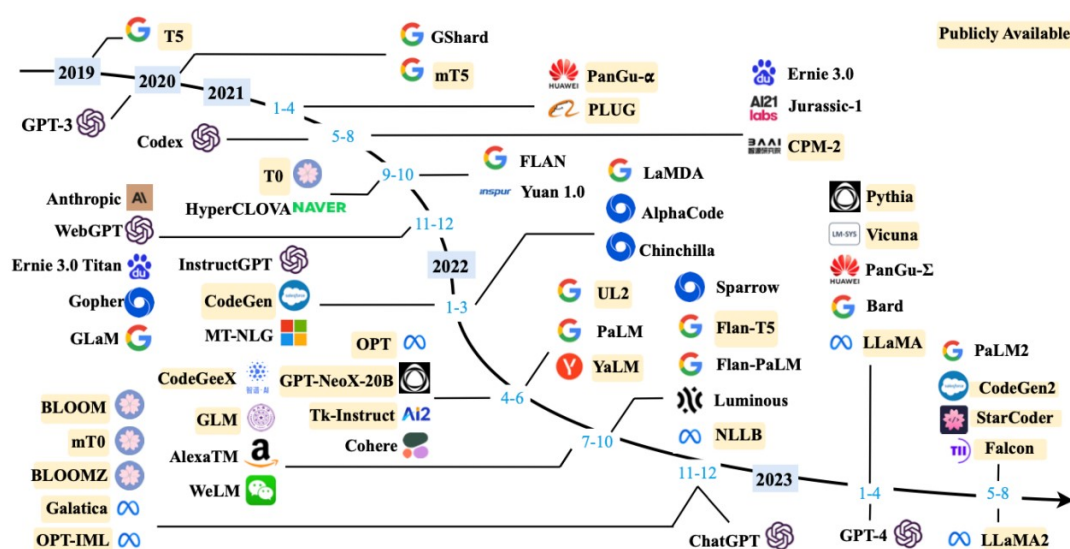


Figure 3: A diagram showing the evolution of publicly available LLMs. Source: Zhao et al. [364].

主要亮点

- 详细探讨了大型语言模型（LLMs）的发展历程、扩展规律和架构策略，揭示了数据和计算资源如何推动其性能提升。
- 分析了 LLMs 在多个领域的应用，包括医疗、金融、教育和法律，展示了其在解决特定领域问题中的适应性和潜力。

- 提供了 LLMs 的能力和局限性，包括如何进行常识推理、代码生成和数学运算等，并探讨了如何利用这些模型执行复杂任务。
- 研究了如何通过 Chain-of-Thought (CoT) 和 Plan-of-Thought (PoT) 方法提升 LLMs 的推理能力，探讨了预训练数据如何影响这些能力的出现。
- 介绍了 LLM 与外部系统的集成框架，分析了这些模型如何通过与外部系统的协同工作处理复杂任务。
- 强调了 LLMs 发展中的伦理问题、技术挑战及其未来应用的潜力，提倡负责任地发展和应用这些技术。

相关链接

Paper: <https://arxiv.org/pdf/2501.04040>

● 论文十二

Test-time Computing: from System-1 Thinking to System-2 Thinking

日期

2025-1-5

Test-time Computing: from System-1 Thinking to System-2 Thinking

Yixin Ji¹, Juntao Li^{1*}, Hai Ye², Kaixin Wu³, Jia Xu³, Linjian Mo³, Min Zhang¹

¹School of Computer Science and Technology, Soochow University

²Department of Computer Science, National University of Singapore

³Ant Group

jiyixin169@gmail.com; {ljt, minzhang}@suda.edu.cn

yehai@comp.nus.edu.sg;

{daniel.wkx, steve.xuj, linyi01}@antgroup.com

内容摘要

本文对“测试时计算” (Test-time Computing) 进行了全面研究调查, 详细探讨了从系统 1 思考 (System-1 Thinking) 到系统 2 思考 (System-2 Thinking) 的转变过程, 并分析了不同模型如何通过测试时计算来提升推理能力。系统 1 模型依赖于快速直觉的推理, 而系统 2 模型则通过更深度的思考和推理解决复杂问题。测试

时计算的应用从系统 1 的分布转移到系统 2，通过多种策略（如重复采样、自我修正和树搜索）逐步增强模型的推理能力。本文对这些策略进行了深入分析，并提出了未来的研究方向，尤其是在提高推理深度和计算效率方面。

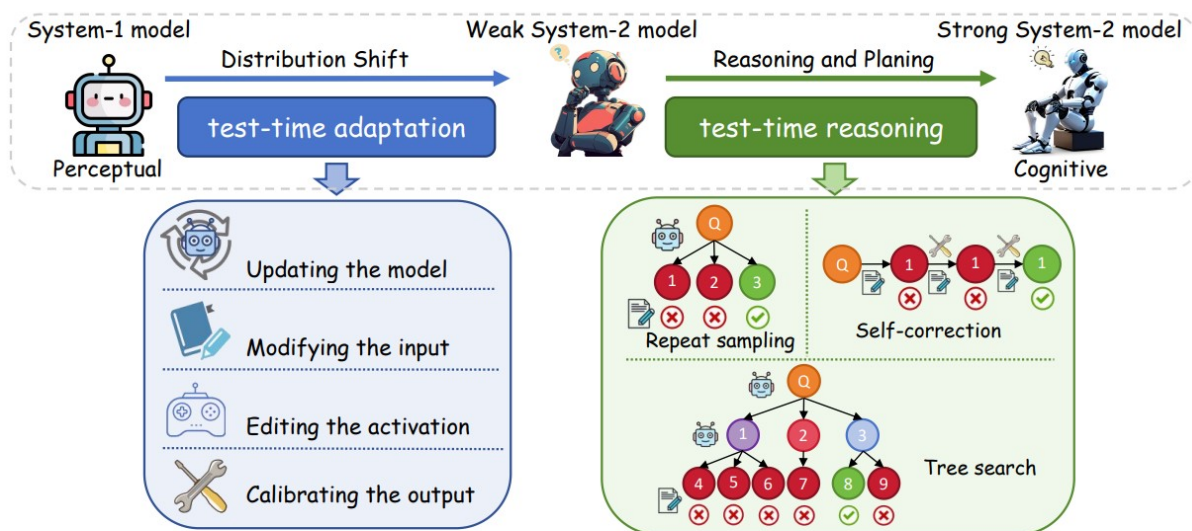


Figure 1: Illustration of test-time computing in the System-1 and System-2 model.

主要亮点

- 测试时计算的核心概念：探讨了测试时计算在系统 1 模型中如何通过参数更新、输入修改、表示编辑和输出校准提高模型的鲁棒性和泛化能力。
- 从系统 1 到系统 2 的过渡：分析了如何通过测试时计算策略（如重复采样、自我修正、树搜索）将系统 1 模型逐步转变为更强大的系统 2 模型，从而增强推理能力。
- 多种推理方法的比较：介绍了测试时推理的三种核心策略（重复采样、自我修正和树搜索），并探讨了这些方法在复杂任务中的有效性。
- 未来研究方向：提出了将多种测试时计算策略结合应用的可能性，以及如何结合推理能力与高效计算达到系统 2 推理的最佳效果。

相关链接

Paper: <https://arxiv.org/pdf/2501.02497>

Github: https://github.com/Dereck0602/Awesome_Test_Time_LLMs

● 论文十三

From Generation to Judgment: Opportunities and Challenges of LLM-as-a-judge

日期

2025-1-6



From Generation to Judgment: Opportunities and Challenges of LLM-as-a-judge

Dawei Li^{*}, Bohan Jiang^{*}, Liangjie Huang^{*}, Alimohammad Beigi^{*}, Chengshuai Zhao^{*},
Zhen Tan^{*}, Amrita Bhattacharjee^{*}, Yuxuan Jiang[†], Canyu Chen[‡], Tianhao Wu^{*},
Kai Shu[§], Lu Cheng^{*}, Huan Liu^{*}

^{*}Arizona State University, [†]University of Illinois Chicago,

[‡]University of Maryland, Baltimore County, [§]Illinois Institute of Technology,

[§]University of California, Berkeley, [§]Emory University

内容摘要

本文深入探讨了“LLM-as-a-judge”（大语言模型作为评判者）的概念，分析了这一新兴领域的机会与挑战。传统的评估方法常常无法精准捕捉细微的属性，而 LLM 作为评判者利用大型语言模型的强大推理能力，可以对不同任务进行打分、排序或选择，成为更精确且灵活的评判工具。本文提供了一个全面的分类框架，从评判的内容、方法和应用场景等三个维度进行详细分析，并提出了 LLM 作为评判者的评估基准。最后，文中还总结了该领域面临的主要挑战，如评判偏差、动态复杂判断以及人类与 LLM 共同评判等问题，并提出了未来的研究方向。



Figure 3: LLMs are capable of judging various attributes.

主要亮点

- 提出了“LLM-as-a-judge”概念，并通过分类框架从评判内容、方法和应用等三个维度深入分析了这一新兴研究领域。
- 通过总结现有的研究工作，探讨了 LLM 作为评判者在多个任务中的应用，如生成任务、推理任务和多模态任务。
- 提供了多种评判方法的详细说明，包括点对点评判、排名评判和选择评判等，并展示了这些方法在不同任务中的实际应用。
- 提出了 LLM 作为评判者的评估基准，并总结了当前的技术挑战，如如何处理评判偏差、如何应对动态和复杂的判断任务。
- 强调了未来研究的潜力，包括自我评判、偏好学习以及如何处理多方协作等新兴方向。

相关链接

Paper: <https://arxiv.org/pdf/2411.16594>

Project: <https://llm-as-a-judge.github.io/>

Github: <https://github.com/llm-as-a-judge/Awesome-LLM-as-a-judge>

● 论文十四

Large language models for artificial general intelligence (AGI): A survey of foundational principles and approaches

日期

2025-1-6

Large language models for artificial general intelligence (AGI): A survey of foundational principles and approaches

Alhassan Mumuni^{1*} and Fuseini Mumuni²

内容摘要

本文深入探讨了大型语言模型（LLMs）如何在实现人工通用智能（AGI）过程中发挥作用，提出了实现 AGI 的四个基础问题：具身化（embodiment）、符号基础（symbol grounding）、因果推理（causality）和记忆机制（memory）。作者认为，尽管现有的 LLMs 已能够解决多种复杂的开放性任务，但其能力仍然局限，尚未达到人类水平的普遍智能。要实现 AGI，LLMs 需要解决涉及这些基础问题的挑战，这些问题与人类认知过程密切相关。本文回顾了如何将这些认知基础应用于 LLMs，并探讨了如何通过多模态学习和其他方法来提升模型的能力，最终朝着具备自主性、推理和适应能力的智能体方向发展。

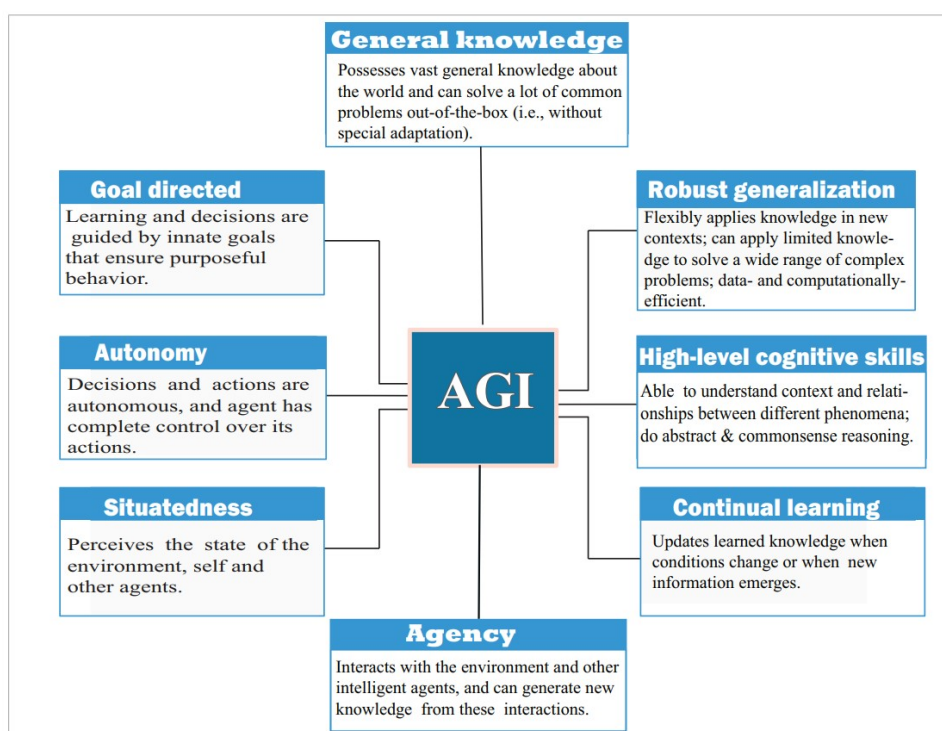


Figure 1. Some of the most important features of artificial general intelligence (AGI) systems. These features give AGI systems vast cognitive capabilities despite the models' limited knowledge and the need to, for the sake of conserving energy and time, take shortcuts in cognitive information processing.

主要亮点

- 提出了 AGI 实现的四个基础原则：具身化、符号基础、因果推理和记忆机制，强调这些概念对于实现人类级别智能至关重要。
- 讨论了大型语言模型如何通过整合多模态数据和学习多维度的世界表示来推动智能体的发展，尤其是在复杂推理和跨领域任务中的应用。
- 强调了具身化（即实体的感知与行动）在智能体实现自我意识、情境意识和目标意识中的关键作用。
- 提出了符号基础在将抽象符号与现实世界的实际对象和概念联系起来方面的作用，解决了模型的抽象理解问题。
- 探讨了因果推理如何帮助智能体理解并预测事件之间的因果关系，提升其处理复杂任务的能力。
- 讨论了如何通过记忆机制帮助模型积累和利用过去的知识，以提高其长期学习能力和适应性。

相关链接

Paper: <https://arxiv.org/pdf/2501.03151>

● 论文十五

Dolphin: Closed-loop Open-ended Auto-research through Thinking, Practice, and Feedback

日期

2025-1-7



DOLPHIN: Closed-loop Open-ended Auto-research through Thinking, Practice, and Feedback

Jiakang Yuan^{1,2}, Xiangchao Yan², Botian Shi², Tao Chen^{1,✉}, Wanli Ouyang²
Bo Zhang^{2,†,✉}, Lei Bai^{2,✉}, Yu Qiao², Bowen Zhou²

¹Fudan University, ²Shanghai Artificial Intelligence Laboratory

内容摘要

本文提出了 DOLPHIN，首个闭环的开放式自动化科学研究框架。DOLPHIN 旨在通过自动生成研究想法、执行实验和基于实验结果生成反馈来实现自动化科研。该框架包含三个主要步骤：1) 生成新想法，2) 进行实验验证，3) 根据实验结果提供反馈并指导下一轮的想法生成。为了提高自动研究的效率，DOLPHIN 通过任务属性引导论文排名和错误回溯引导调试过程，优化了想法质量和代码执行成功率。实验结果表明，DOLPHIN 在多个基准数据集上生成的想法与现有的最先进方法相媲美，且在某些任务上表现出色，能够自动完成实验并持续改进研究效果。

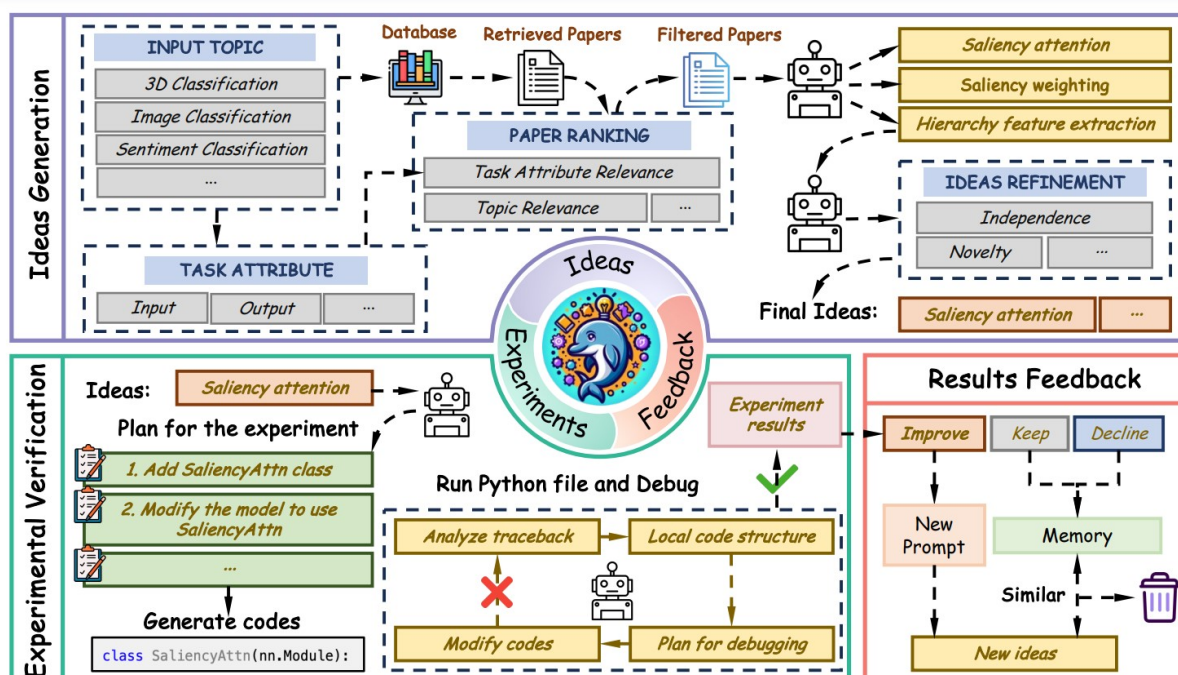


Figure 2. The overview of DOLPHIN, an closed-loop open-ended auto-research framework. DOLPHIN first generates a set of ideas based on the retrieved papers. After filtering ideas, experimental plans will be generated for these filtered ideas. Then, codes can be generated and debugged using the proposed error-traceback-guided debugging process. Finally, the results of successfully executed experiments will be auto-analyzed and reflected into the next round of ideas generation.

主要亮点

- 提出了 DOLPHIN 框架，首次实现了闭环的自动化科学研究过程，涵盖了从想法生成、实验验证到结果反馈的完整科研循环。
- 提出了一种任务属性引导的论文排名和基于错误回溯的调试方法，显著提高了生成想法的质量和代码执行的成功率。
- 通过实验验证，DOLPHIN 能够自动生成与现有最先进方法相媲美的研究想法，并且在多个任务（如图像分类、3D 点分类、情感分类）上取得了较大的性能提升。
- 强调了 DOLPHIN 在生成高质量研究想法方面的创新性，尤其是在 3D 点分类任务中，通过自动化研究方法达到了与人类设计的方法相当的效果。
- 在反馈机制上，DOLPHIN 能够通过分析实验结果，不断优化生成的想法，有效避免重复验证无效的想法，提高了研究效率。

相关链接

Paper: <https://arxiv.org/pdf/2501.03916>

Project: <https://alpha-innovator.github.io/Dolphin-project-page/>

Code: <https://github.com/Alpha-Innovator/Dolphin>

● 论文十六

Accurate predictions on small data with a tabular foundation model

日期

2025-1-8

Article

Accurate predictions on small data with a tabular foundation model

<https://doi.org/10.1038/s41586-024-08328-6>

Received: 17 May 2024

Noah Hollmann^{1,2,3,7}✉, Samuel Müller^{1,7}✉, Lennart Purucker¹, Arjun Krishnakumar¹,
Max Körfer¹, Shi Bin Hoo¹, Robin Tibor Schirrmeyer^{4,5} & Frank Hutter^{1,3,6}✉

内容摘要

本文介绍了 TabPFN，一种用于小型到中型表格数据的基础模型，能够在表格数据的分类和回归任务中显著优于传统方法。TabPFN 采用了生成模型，通过在数百万个合成数据集上进行训练来自动学习有效的预测算法，并应用于未见过的真实数据集，避免了手动设计和调试的需要。该模型基于变压器（Transformer）架构，利用“上下文学习”（ICL）机制，通过一次前向传递实现对标签的预测，从而大幅提高了处理速度。TabPFN 的性能在多个标准数据集上均超越了包括 XGBoost、CatBoost 等在内的当前最强基线方法，同时支持数据生成、密度估计和可重用嵌入的学习。

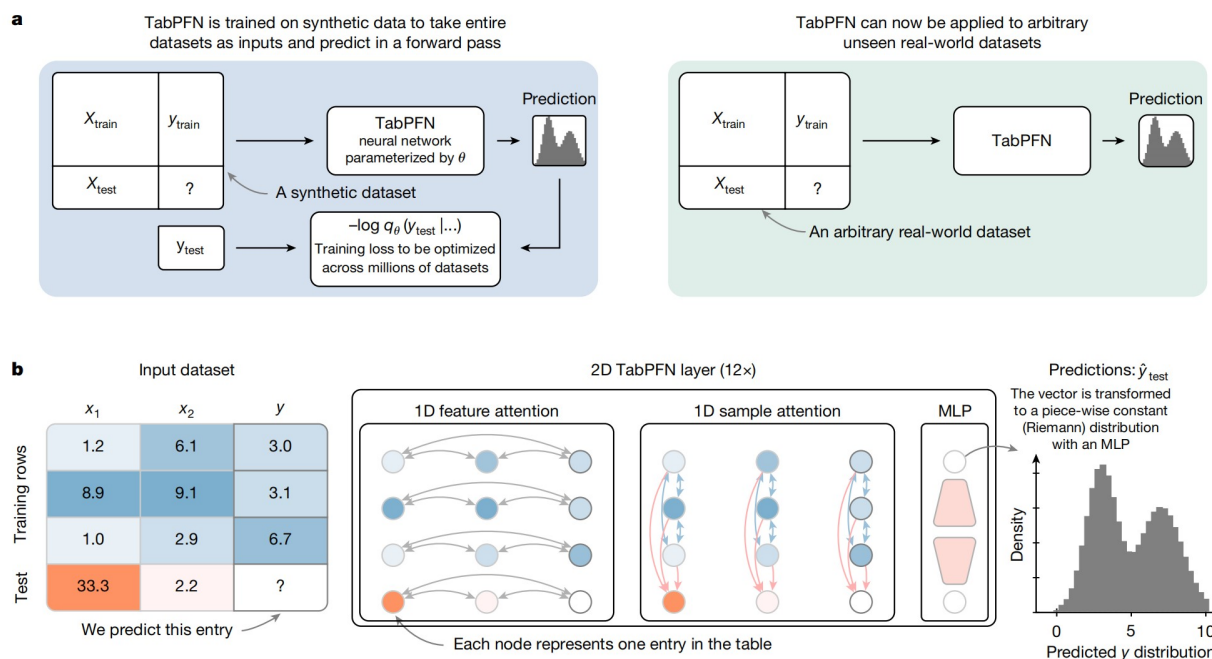


Fig. 1 | Overview of the proposed method. **a**, The high-level overview of TabPFN pre-training and usage. **b**, The TabPFN architecture. We train a model to solve more than 100 million synthetic tasks. Our architecture is an adaptation of the

standard transformer encoder that is adapted for the two-dimensional data encountered in tables.

主要亮点

- 提出了 TabPFN 模型，通过生成合成数据集进行训练，能够自动学习处理小型表格数据的有效预测算法，显著提升了表格数据处理效率。
- 在表格数据的分类和回归任务中，TabPFN 在多个基准测试中超越了现有的主流方法，如 XGBoost 和 CatBoost，且在计算速度上实现了显著提升。
- TabPFN 支持在未知数据集上的推理，能够在一次前向传递中完成训练和预测，大幅提高了推理效率。
- 利用上下文学习（ICL）机制，TabPFN 不仅能够进行高效的预测，还能够进行数据生成、密度估计以及学习可重用的嵌入，扩展了其在实际应用中的适用性。
- TabPFN 的设计特别适用于小到中型数据集，处理最大样本量为 10,000 的表格数据，并能够在有限的计算资源下快速执行。

相关链接

Paper: <https://www.nature.com/articles/s41586-024-08328-6>

● 论文十七

On Computational Limits and Provably Efficient Criteria of Visual Autoregressive Models: A Fine-Grained Complexity Analysis

日期

2025-1-8

On Computational Limits and Provably Efficient Criteria of Visual Autoregressive Models: A Fine-Grained Complexity Analysis

Yekun Ke* Xiaoyu Li† Yingyu Liang‡ Zhizhou Sha§ Zhenmei Shi¶
Zhao Song^{||}

内容摘要

本文研究了视觉自回归 (VAR) 模型的计算限制及其高效算法标准。VAR 模型通过一种从粗到细的“下一个尺度预测”方法推进图像生成,突破了传统方法的计算瓶颈。然而,现有的 VAR 算法的计算复杂度高达 $O(n^4)$, 存在明显的计算效率问题。本文分析了 VAR 模型的计算限制,并提出了在某些条件下可以实现亚二次 (sub-quadratic) 时间复杂度的算法。具体来说,本文基于强指数时间假设 (SETH) 确立了输入矩阵的范数阈值,并在此阈值下证明了 VAR 模型的计算难度。研究还展示了低秩近似的有效构造,并提出了两种重要的结果: 1) 当输入矩阵范数 R 小于某一阈值时,可以设计出 $O(n^2+o(1))$ 的算法来近似 VAR 模型; 2) 当范数大于阈值时,无法在亚四次时间内近似 VAR 模型。

Can we perform the computations of VAR models faster than $O(n^4)$ time?

We answer this question affirmatively and summarize our contributions as follows.

- **Computational Limits:** We analyze the computation of the VAR models under the Strong Exponential Time Hypothesis. Let R represent the upper bound of the elements in the input matrices used for attention calculations in VAR models. We establish an upper bound criterion $R^* = \Theta(\sqrt{\log n})$. Crucially, only when R is below this threshold, one can compute VAR models in $O(n^{4-\Omega(1)})$ time (truly sub-quartic time).
- **Provably Efficient Criteria:** We further show that when $R = o(\sqrt{\log n})$, it becomes possible to design an algorithm that approximates the VAR model in almost quadratic time, specifically $O(n^{2+o(1)})$.

主要亮点

- 提出了 VAR 模型计算的计算复杂度分析，并在理论上界定了其计算极限，阐明了何种情况下 VAR 模型能够达到亚二次时间复杂度。
- 通过强指数时间假设 (SETH)，证明了当输入矩阵的范数超过某一阈值时，无法设计出亚四次时间复杂度的算法。
- 展示了低秩近似的有效构造方法，这为 VAR 模型的计算效率提供了改进的方向。
- 提出了一个新的算法，在特定条件下实现了 $O(n^2 + o(1))$ 的时间复杂度，可以高效地近似 VAR 模型的输出。
- 为 VAR 模型的计算瓶颈提供了理论框架，并为未来高效的图像生成算法设计提供了理论基础。

相关链接

Paper: <https://arxiv.org/pdf/2501.04377>

● 论文十八

rStar-Math: Small LLMs Can Master Math Reasoning with Self-Evolved Deep Thinking

日期

2025-1-8

rStar-Math: Small LLMs Can Master Math Reasoning with Self-Evolved Deep Thinking

Xinyu Guan* Li Lyna Zhang** Yifei Liu
Ning Shang Youran Sun Yi Zhu Fan Yang Mao Yang

Microsoft Research Asia

内容摘要

本文提出了 rStar-Math，一种通过自我演化的深度思维方法，显著提升小型语言模型（SLMs）在数学推理方面的能力。rStar-Math 通过结合蒙特卡罗树搜索（MCTS）和过程奖励模型（PPM），使小型语言模型能够自我生成高质量的训练数据，从而进行复杂的数学推理。该方法通过四轮自我进化，逐步优化推理过程和奖励模型，最终在多个数学基准测试中超越了现有的大型语言模型，如 OpenAI 的 o1 模型。rStar-Math 展示了在解决复杂数学问题（如 MATH、AIME 等竞赛级别问题）方面的强大能力，特别是在推理步骤的验证和提升方面，取得了显著的进展。在 MATH 测试集上，可以把 Qwen2.5-Math-7B 性能从 58.8% 提升到 90.0%，超过 o1-preview。

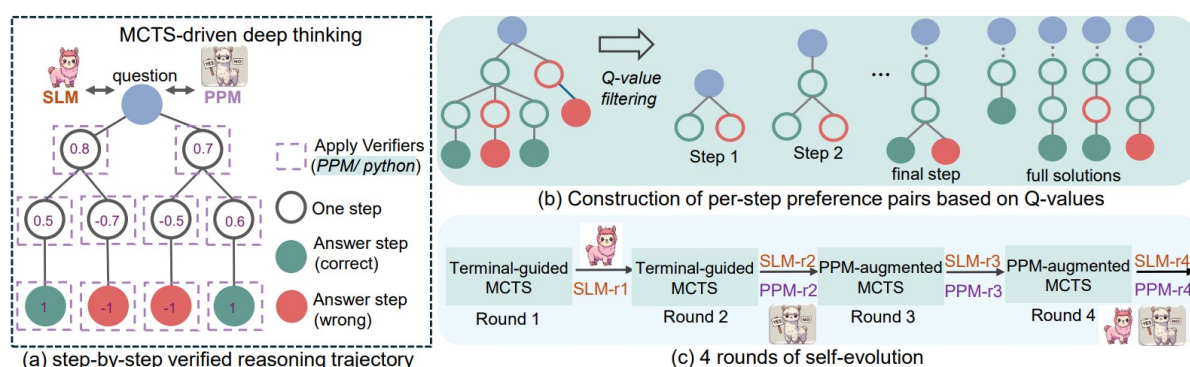


Figure 1: The overview of rStar-Math.

主要亮点

- 提出了 rStar-Math 框架，利用 MCTS 和过程奖励模型（PPM）进行深度思维，通过自我演化过程提升数学推理能力。
- 采用创新的代码增强 CoT 数据合成方法，通过多轮 MCTS 展开生成逐步验证的推理轨迹，保证推理步骤的准确性。
- 提出了新的过程奖励模型训练方法，通过偏好对比（PPM）避免了传统方法中依赖于精确步骤级奖励的挑战，提升了训练效果。

- 通过四轮自我进化，显著提升了小型模型（如 Qwen2.5-Math-7B）的推理能力，在 MATH 基准测试中将其准确率从 58.8%提升到 90.0%。
- 该方法不仅适用于数学推理，还具有广泛的推广潜力，可以应用于代码推理和常识推理等其他领域。

相关链接

Paper: <https://arxiv.org/pdf/2501.04519>

Github: <https://github.com/microsoft/rStar>

● 论文十九

The Future of AI: Exploring the Potential of Large Concept Models

日期

2025-1-8

The Future of AI Exploring the Potential of Large Concept Models

Hussain Ahmad^{†*}
The University of Adelaide, Australia
hussain.ahmad@adelaide.edu.au

Diksha Goel^{*}
CSIRO's Data61, Australia
diksha.goel@data61.csiro.au

内容摘要

本文探讨了大型概念模型（Large Concept Models, LCMs）的潜力，这是一种相较于传统大型语言模型（Large Language Models, LLMs）更加先进的人工智能框架。LCMs 通过将处理单位从单一的词或子词转变为整个语义单元（即概念），显著改善了长文本上下文理解、多模态推理和跨语言处理的能力。文章系统性地分析了 LCMs 的独特特性、应用场景以及对未来研究和实践的意义，并提出了优化 LCMs 开发和应用的建议。

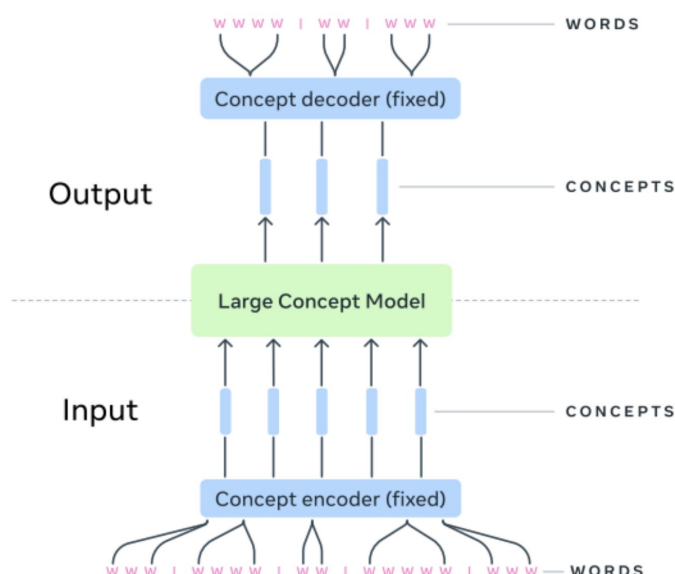


Fig. 2: Fundamental architecture of Large Concept Model [17].

主要亮点

- 引入概念级推理模型：LCMs 将处理单位从传统 LLMs 的词级别提升到概念级别（句子或完整思想），实现了更高层次的语义推理能力，增强了上下文一致性和结构性输出
- 多模态及多语言支持：通过 SONAR 嵌入技术，LCMs 能够在 200 多种语言和多个模态（如文本、语音）之间实现无缝操作，不需要额外的训练即可处理低资源语言或跨模态任务。
- 高效处理长上下文：LCMs 通过减少处理单位（从数千词到少量概念）降低了计算复杂度，能够高效处理长文档，同时保持文本的逻辑连贯性和语义一致性。
- 广泛的实际应用场景：LCMs 在多领域展现了强大的应用潜力，包括网络安全、医疗健康、教育、科学研究、法律分析、个性化推荐等，大幅提升了任务效率和决策准确性。
- 为研究与实践提供指导：文章提出了一系列研究方向和实践建议，包括优化概念嵌入空间、改进量化策略、增强跨语言和多模态通用性，为 LCMs 的进一步发展提供了清晰的路线图。

相关链接

Paper: <https://arxiv.org/pdf/2501.05487>

Github: https://github.com/facebookresearch/large_concept_model

● 论文二十

Toward expert-level medical question answering with large language models

日期

2025-1-8

nature medicine

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [nature medicine](#) > [articles](#) > [article](#)

Article | [Open access](#) | Published: 08 January 2025

Toward expert-level medical question answering with large language models

[Karan Singhal](#), [Tao Tu](#), [Juraj Gottweis](#), [Rory Sayres](#), [Ellery Wulczyn](#), [Mohamed Amin](#), [Le Hou](#), [Kevin Clark](#), [Stephen R. Pfohl](#), [Heather Cole-Lewis](#), [Darlene Neal](#), [Qazi Mamunur Rashid](#), [Mike Schaekermann](#), [Amy Wang](#), [Dev Dash](#), [Jonathan H. Chen](#), [Nigam H. Shah](#), [Sami Lachgar](#), [Philip Andrew Mansfield](#), [Sushant Prakash](#), [Bradley Green](#), [Ewa Dominowska](#), [Blaise Agüera y Arcas](#), [Nenad Tomašev](#), ... [Vivek Natarajan](#) ✉ [+ Show authors](#)

内容摘要

本文提出了 Med-PaLM 2，一种为医学问题解答优化的大型语言模型（LLM）。通过改进的基础语言模型（PaLM 2）、医学领域的微调、集合优化（ensemble refinement）和检索链（chain of retrieval）等策略，Med-PaLM 2 在多项医学基准测试上取得了显著进展，例如 MedQA、MedMCQA、PubMedQA 与 MMLU 临床主题数据集，其性能超过了此前的 Med-PaLM 和其他主流 LLM。此外，Med-PaLM 2 在长文本医学问答中表现优异，医生评估显示其答案在多个临床相关维度上优于普通医生生成的答案，并与专家医生答案相当。通过引入对抗性数据集和真实临床问题数据集，进一步验证了模型在安全性、事实性和临床实用性方面的可靠性，展示了其在医学实际应用中的潜力。

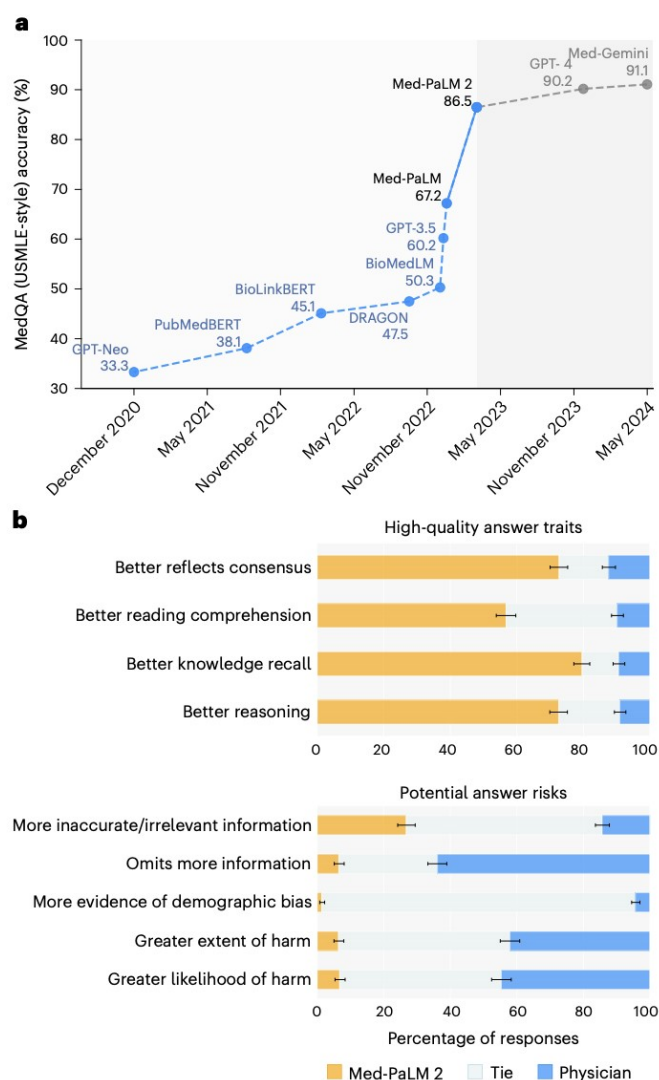


Fig. 1 | Med-PaLM 2 performance on MultiMedQA. a, Med-PaLM 2 achieved an accuracy of 86.5% on USMLE-style questions in the MedQA dataset. The shaded region highlights the reported performance of models developed after Med-PaLM 2. **b**, In a pairwise ranking study on $n = 1,066$ consumer medical questions, Med-PaLM 2 answers were preferred over physician answers by a panel of physicians across eight of nine axes in our evaluation framework. Stacked bars represent proportions of answers for which physician raters preferred Med-PaLM 2 answers (orange), answers generated by other physicians (blue) or ties (light blue). Error bars reflect 95% confidence intervals of the overall preference rates for physician and Med-PaLM 2 answers, as determined by clustered bootstrapping computed over all 1,066 paired ratings.

主要亮点

- 提出 Med-PaLM 2 模型：基于改进的 PaLM 2 语言模型，通过医学领域的微调和新策略优化，Med-PaLM 2 在多项医学基准测试上取得了突破性成绩，MedQA 得分高达 86.5%，较前代模型提升 19% 以上。

- 引入集合优化 (ensemble refinement)：作为一种新的提示策略，集合优化通过综合多个生成路径提升模型的推理能力，显著改善了多项选择题的解答效果。
- 检索链策略提升事实性：提出检索链方法，通过分步骤验证信息来源，增强模型在复杂医学问题上的事实性和安全性，为真实临床问题提供更可靠的答案。
- 人类评估显示优异表现：在多维度评估中，Med-PaLM 2 的长文本医学答案在 8 个临床维度上优于普通医生生成的答案，包括事实性、推理能力和低风险性等，同时在 90.6% 的答案中显示出低伤害风险。
- 实际应用潜力：在真实临床问题上，Med-PaLM 2 生成的答案被专家医生在 65% 的情况下优于普通医生答案，显示出未来在医疗辅助决策中的应用前景，同时其安全性被评估为与医生答案相当。

相关链接

Paper: <https://www.nature.com/articles/s41591-024-03423-7>

● 论文二十一

LlamaV-o1: Rethinking Step-by-step Visual Reasoning in LLMs

日期

2025-1-10



LlamaV-o1: Rethinking Step-by-step Visual Reasoning in LLMs

Omkar Thawakar^{1*} Dinura Dissanayake^{1*} Ketan More^{1*} Ritesh Thawkar^{1*}
Ahmed Heakl^{1*} Noor Ahsan^{1*} Yuhao Li^{1*} Mohammed Zumri^{1*} Jean Lahoud^{1*}
Rao Muhammad Anwer¹ Hisham Cholakkal¹ Ivan Laptev¹ Mubarak Shah²
Fahad Shahbaz Khan^{1,3} Salman Khan^{1,4}

¹Mohamed bin Zayed University of AI, ²University of Central Florida,
³Linköping University, ⁴Australian National University

内容摘要

本文提出了一种名为 LlamaV-o1 的新型多模态视觉推理模型，重点解决复杂的多步视觉推理任务。为了推动大语言模型（LLMs）在视觉推理中的逐步推理能力，研究人员构建了一个全面的评估框架，并提出了三个主要贡献：一个新的视觉推理数据集（VRC-Bench）、一种评估多步推理质量的新指标，以及一种通过课程学习逐步训练的多模态推理模型 LlamaV-o1。实验结果表明，LlamaV-o1 在多个基准数据集上优于现有的开源模型，并在一些任务上达到了与闭源模型相当的性能，同时推理速度显著提升。

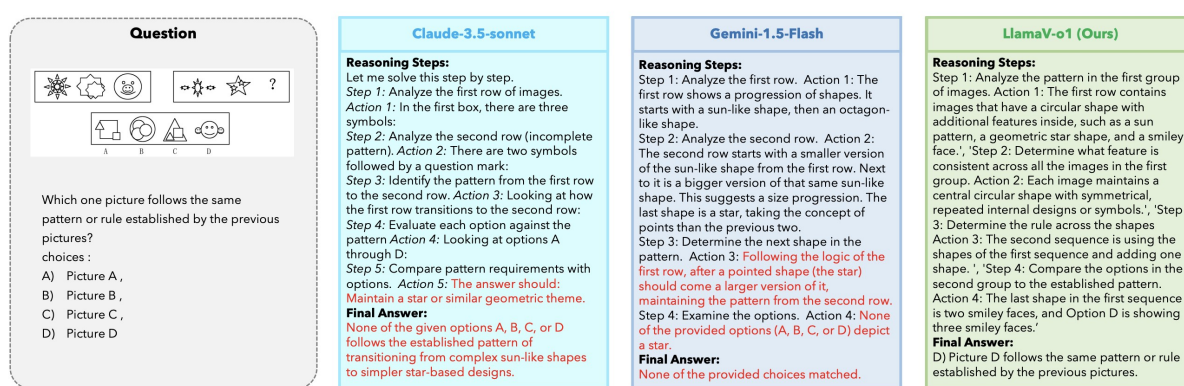


Figure 1: Comparison of the reasoning abilities of our model (LlamaV-o1) with closed-source Gemini-1.5-Flash and Claude-3.5-Sonnet on an example in pattern recognition task from our proposed VRC-Bench. While Claude-3.5-Sonnet concludes "none of the options," its reasoning steps lack full alignment with the observed logic (highlighted in red). Gemini-1.5-Flash demonstrates weaker reasoning with less logical coherence details (highlighted in red). Our LlamaV-o1 provides better and more systematic reasoning, identifying that option D follows the established pattern, thereby showcasing its logical reasoning capability. Additional results are presented in Fig. 5.

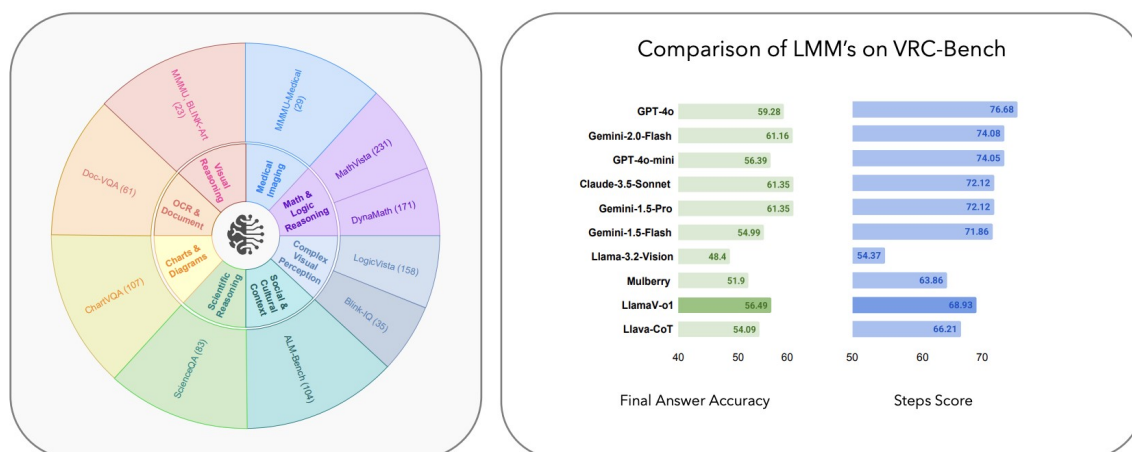


Figure 3: The figure illustrates our comprehensive benchmark structure and comparative performance of LMMs on the proposed ReasoningChain-Bench. (Left) The dataset spans multiple domains, including carefully selected samples for mathematical and logical reasoning (e.g., MathVista [38] with 231 samples and LogicVista with 158 samples), scientific reasoning (e.g., Science-QA [40] with 83 samples), and visual perception (e.g., Blink-IQ-Test [15] with 35 samples). Additionally, it includes specialized areas such as medical imaging (e.g., MMMU-Medical [72] with 29 samples), cultural and social understanding (e.g., ALM-Bench [57] with 104 samples), and document understanding through OCR (e.g., Doc-VQA [46] with 61 samples). By integrating tasks like chart and diagram comprehension (e.g., Chart-VQA [44] with 107 samples), our dataset not only covers a broad spectrum of real-world applications but also expand LMM's ability to reason, perceive, and interpret complex multimodal information. (Right) The bar chart compares various SoTA reasoning models on the VRC-Bench, highlighting both final answer accuracy and step-by-step reasoning scores. The models evaluated for complex reasoning tasks include GPT-4o, Gemini-2.0-Flash, Claude-3.5-Sonnet, and Llava-CoT. Our benchmark evaluates models not only on their ability to generate accurate final answers but also on the coherence and logical flow of their reasoning steps. Our approach, LlamaV-o1, outperforms GPT-4o-mini, Gemini-1.5-Flash and Llava-CoT in the VRC-Bench, achieving superior results in final answer accuracy across complex multimodal reasoning tasks.

主要亮点

- **提出了新的视觉推理基准 VRC-Bench:** VRC-Bench 是一个专门设计用于评估多步视觉推理任务的综合性基准数据集，涵盖了八个类别（视觉推理、数学与逻辑推理、社会文化背景、医学影像、图表理解、OCR 文档理解、复杂视觉感知、科学推理），包含 1000 多个样本和超过 4173 步的推理链。
- **设计了一种新型评估指标:** 提出了一个能够以单步粒度评估推理质量的指标，强调逻辑连贯性与正确性，克服了传统只关注最终任务准确率的局限性。
- **开发了 LlamaV-o1 多模态推理模型:** 采用课程学习方法进行训练，从简单任务逐步过渡到复杂多步推理任务，显著提升了模型的逻辑推理能力和解释性。

- **引入高效的推理优化方法**：结合了多步课程学习和 Beam Search 技术，实现了推理效率的显著提高。相比现有方法（如 Llava-CoT），LlamaV-o1 在推理速度上快了 5 倍，同时保持高准确率。
- **优于现有模型的性能表现**：在多个基准数据集（如 MathVista、AI2D、HallusionBench 等）和新提出的 VRC-Bench 上，LlamaV-o1 的多步推理得分和最终答案准确率均领先现有开源模型（如 Llava-CoT），并在一些复杂任务中接近闭源模型（如 GPT-4o）。

相关链接

Paper: <https://arxiv.org/pdf/2501.06186>

Github: <https://github.com/mbzuai-oryx/LlamaV-o1>

Model: <https://github.com/mbzuai-oryx/LlamaV-o1>

Bench: <https://huggingface.co/datasets/omkarthawakar/VRC-Bench>

● 论文二十二

The Lessons of Developing Process Reward Models in Mathematical Reasoning

日期

2025-1-13

The Lessons of Developing Process Reward Models in Mathematical Reasoning

Zhenru Zhang Chujie Zheng Yangzhen Wu Beichen Zhang Runji Lin
Bowen Yu* Dayiheng Liu* Jingren Zhou Junyang Lin*
Qwen Team, Alibaba Group

内容摘要

本文研究了数学推理中的过程奖励模型（Process Reward Models, PRMs），旨在通过对推理过程的监督提升大语言模型（LLMs）的推理能力和可靠性。作者分析了现有 PRM 开发中基于蒙特卡洛（MC）估计、

LLM-as-a-judge (LLM 作为评判者) 和人工标注的不同数据构建方法, 发现 MC 估计存在显著局限性, 导致模型性能和泛化能力不足。此外, 论文还探讨了 Best-of-N (BoN) 评估策略的偏差, 提出了结合 MC 估计与 LLM-as-a-judge 的共识过滤机制, 显著改进了数据效率和模型性能。通过实验, 作者验证了新提出的 PRMs 在响应级别和步骤级别评测中的优越性, 并开源了相关模型, 为未来的 PRM 研究提供了实践指导。

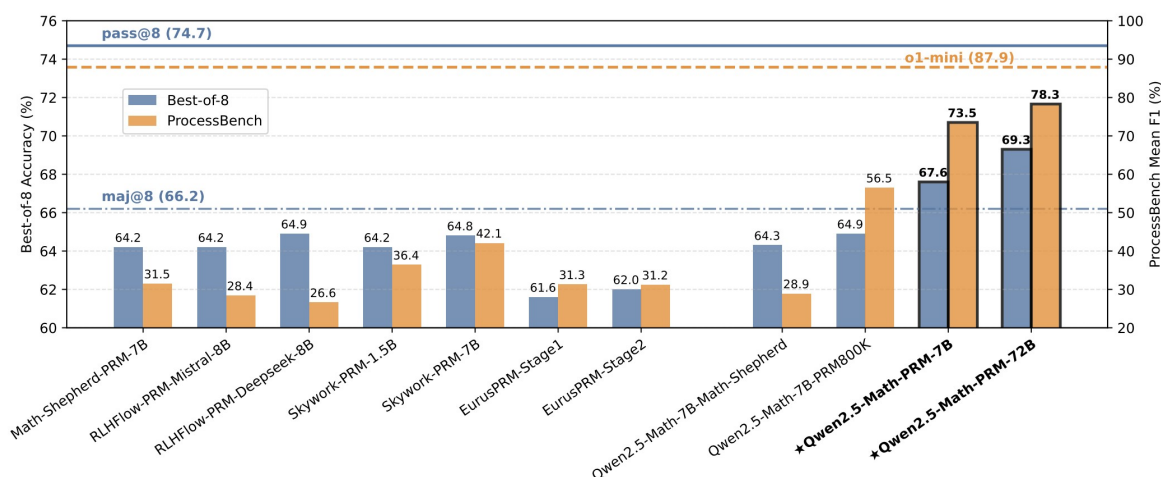


Figure 1: Overview of evaluation results on the Best-of-8 strategy of the policy model Qwen2.5-Math-7B-Instruct and the benchmark PROCESSBENCH (Zheng et al., 2024) across multiple PRMs (see Table 6 and Table 7 for details).

主要亮点

- 揭示 MC 估计的局限性：通过实验表明，基于 MC 估计的数据构建在错误定位和泛化能力上显著劣于 LLM-as-a-judge 和人工标注方法。
- 提出共识过滤机制：将 MC 估计与 LLM-as-a-judge 相结合，通过共识过滤筛选高质量数据，显著提升了 PRM 的训练数据效率和模型表现。
- 优化评估方法：指出 BoN 评估存在显著偏差，可能导致 PRMs 从过程验证转向结果验证，并建议结合步骤级别评测（如 PROCESSBENCH）弥补 BoN 的不足。
- 验证硬标签优于软标签：通过实验发现，在 MC 估计中使用硬标签训练 PRMs 性能更优，且共识过滤进一步增强了硬标签的效果。

- 发布最先进的 PRMs：开源了两个新 PRMs (Qwen2.5-Math-PRM-7B 和 Qwen2.5-Math-PRM-72B) ，在 BoN 评估和 PROCESSBENCH 上均显著优于现有开源模型，为 PRM 开发提供了新的标杆。

相关链接

Paper: <https://arxiv.org/pdf/2501.07301>

Model: <https://huggingface.co/Qwen/Qwen2.5-Math-PRM-7B>;
<https://huggingface.co/Qwen/Qwen2.5-Math-PRM-72B>

● 论文二十三

Transformer2: Self-adaptive LLMs

日期

2025-1-14

TRANSFORMER²: SELF-ADAPTIVE LLMs

Qi Sun^{1,2*}, Edoardo Cetin^{1*}, Yujin Tang^{1*}

¹Sakana AI, Japan ²Institute of Science Tokyo, Japan

{qisun, edo, yujintang}@sakana.ai

*Equal contribution

内容摘要

本文提出了 Transformer2，一种全新的自适应大语言模型（LLM）框架，旨在解决传统微调方法计算成本高、任务适应性差的问题。Transformer2 通过引入奇异值微调（SVF）方法，仅调整权重矩阵的奇异值，从而以更少的参数实现高效的任务适应。框架采用两阶段推理机制：第一阶段通过任务识别系统分析任务特性，第二阶段利用预训练的任务“专家”向量动态调整模型权重以适应特定任务。实验表明，Transformer2 在多个任务上超越了现有方法（如 LoRA），在语言建模、推理、视觉-语言任务等领域展现出卓越性能，同时显著降低参数需求，开辟了自适应 LLM 的新方向。

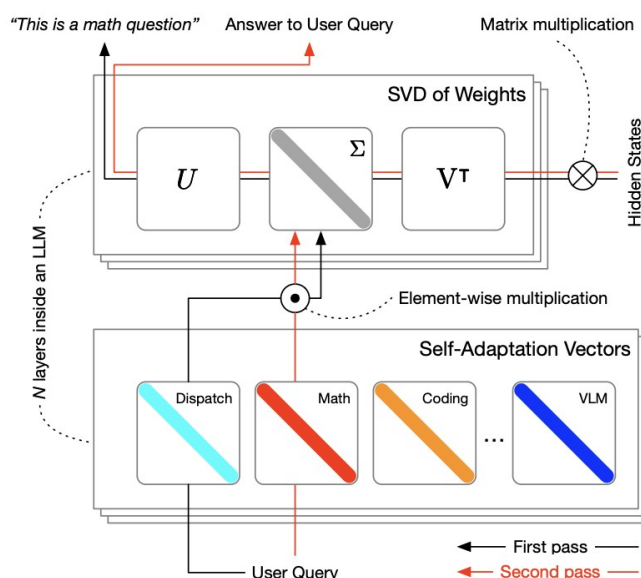


Figure 1: **Overview of Transformer²**. In the training phase, we tune the scales of the singular values of the weight matrices to generate a set of “expert” vectors, each of which specializes in one type of tasks. In the inference phase, a two-pass process is adopted where the first applies the task-specific expert and the second generates the answer.

主要亮点

- 提出 Transformer2 框架：Transformer2 是一种自适应框架，通过动态组合任务“专家”向量，实现 LLM 在推理阶段实时适应新任务，提供了通用的动态适应解决方案。
- 引入奇异值微调（SVF）方法：SVF 通过调整权重矩阵的奇异值，显著减少训练参数，同时具备高效性、可组合性和强正则化能力，与 RL 结合可直接优化任务性能。
- 两阶段推理机制：提出基于任务分类的两阶段推理方法，通过任务识别动态选择或生成适合任务的专家向量，提升了模型在测试时的自适应能力。
- 跨模型适配能力：实验验证了 SVF 专家向量在不同模型间的迁移能力，展现了跨模型通用性，为模型能力复用和民主化提供了新思路。

- 显著提升性能与效率：在多个任务（如数学推理、编程、视觉问答等）中，Transformer2 在性能上显著优于 LoRA 等现有方法，同时训练成本和参数量减少了一个数量级。

相关链接

Paper: <https://arxiv.org/abs/2501.06252>

Github: <https://github.com/SakanaAI/self-adaptive-llms>

● 论文二十四

OmniThink: Expanding Knowledge Boundaries in Machine Writing through Thinking

日期

2025-1-16



OmniThink: Expanding Knowledge Boundaries in Machine Writing through Thinking

Zekun Xi^{1,2}, Wenbiao Yin², Jizhan Fang¹, Jialong Wu², Runnan Fang^{1,2},
Ningyu Zhang^{1,*}, Jiang Yong^{2,*}, Pengjun Xie², Fei Huang², Huajun Chen^{1,3}

¹Zhejiang University

²Tongyi Lab, Alibaba Group

³Zhejiang Key Laboratory of Big Data Intelligent Computing
{xizekun2023, zhangningyu}@zju.edu.cn

内容摘要

本文提出了 OmniThink，一个模仿人类迭代反思与扩展认知过程的机器写作框架。OmniThink 通过连续的“扩展-反思”机制，在检索信息的基础上动态调整检索策略，逐步加深对主题的理解，生成知识密度更高、内容更深入的长文。实验表明，OmniThink 在知识丰富性、一致性、深度和新颖性方面显著优于现有方法，并通过自动评估和人类评估验证了其有效性，为开放领域长文生成提供了新的视角。

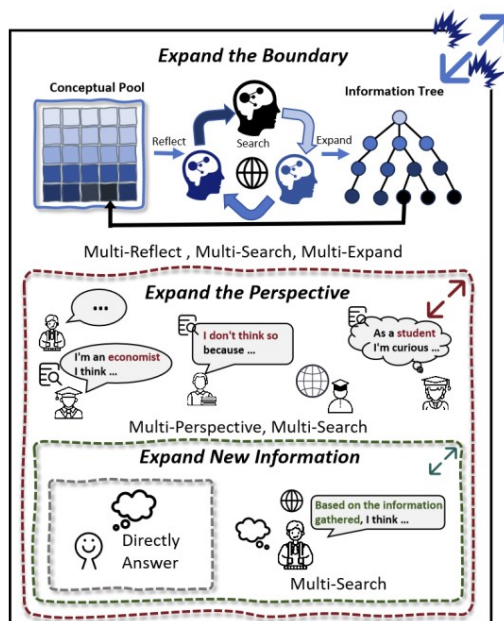


Figure 1: Previous machine writing approaches only expand new information or perspective via RAG and role-playing. OmniThink expands knowledge boundaries through continuous reflection and exploration, attaching knowledge to an information tree and extracting it into a conceptual pool to deepen understanding and uncover more in-depth content.

主要亮点

- 提出迭代扩展与反思机制：OmniThink 引入 “扩展-反思” 循环，通过动态调整检索与生成策略，构建信息树和概念池，逐步加深对主题的理解，提升内容质量。
- 创新知识密度指标：提出了知识密度（Knowledge Density）指标，用于衡量生成文本中有意义内容占比，优化生成文章的冗余问题，提高信息获取效率。
- 显著提升生成文本质量：在 WildSeek 数据集上，OmniThink 在连贯性、广度、深度和新颖性等指标上表现优于现有基线方法，尤其在新颖性和多样性上有显著提升。
- 增强生成内容的结构与逻辑性：通过概念池的设计，OmniThink 生成的文章大纲在结构层次和逻辑一致性上显著优于其他方法，为内容生成提供了更清晰的指导。

- 模型无关性与广泛适用性: OmniThink 框架具有模型无关性, 可与现有框架集成, 同时支持多种语言模型作为生成后端, 展示了其广泛应用潜力。

相关链接

Paper: <https://arxiv.org/pdf/2501.09751>

Project: <https://zjunlp.github.io/project/OmniThink>

Code: <https://github.com/zjunlp/OmniThink>

Demo: <https://www.modelscope.cn/studios/iic/OmniThink>

● 论文二十五

A generative model for inorganic materials design

日期

2025-1-16

A generative model for inorganic materials design

Claudio Zeni ^{1†}, Robert Pinsler ^{1†}, Daniel Zügner ^{2†},
Andrew Fowler ^{1†}, Matthew Horton ^{3†}, Xiang Fu ¹,
Zilong Wang ⁷, Aliaksandra Shysheya¹, Jonathan Crabbé ¹,
Shoko Ueda¹, Roberto Sordillo¹, Lixin Sun ¹, Jake Smith³,
Bichlien Nguyen³, Hannes Schulz², Sarah Lewis ¹,
Chin-Wei Huang⁶, Ziheng Lu ⁴, Yichi Zhou⁵, Han Yang ⁴,
Hongxia Hao ⁴, Jielan Li ⁴, Chunlei Yang⁷, Wenjie Li⁷,
Ryota Tomioka ^{1†}, Tian Xie ^{1*†}

¹Microsoft Research AI for Science, Cambridge, UK.

²Microsoft Research AI for Science, Berlin, Germany.

³Microsoft Research AI for Science, Redmond, USA.

⁴Microsoft Research AI for Science, Shanghai, China.

⁵Microsoft Research AI for Science, Beijing, China.

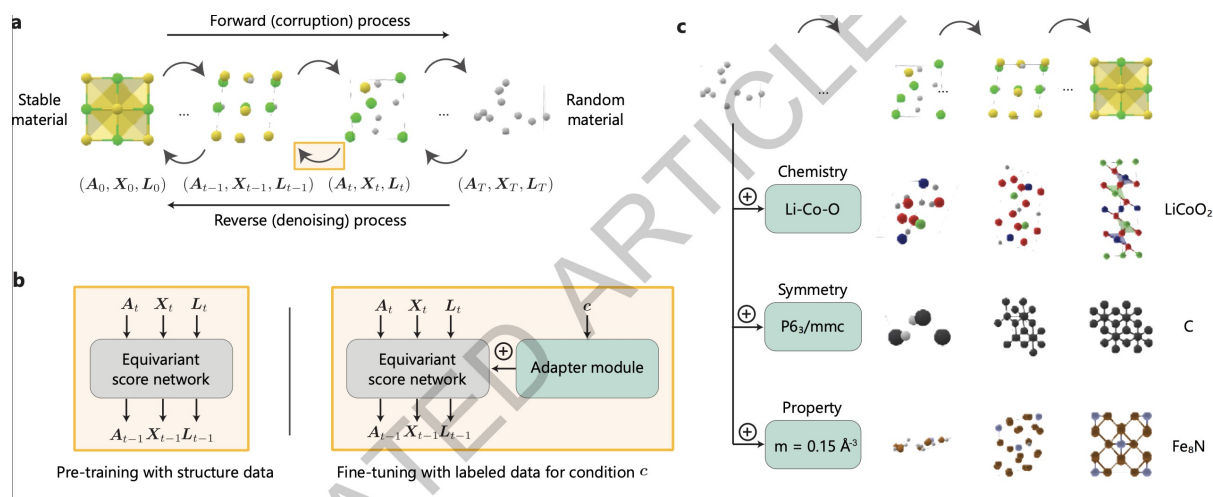
⁶Microsoft Research AI for Science, Amsterdam, Netherlands.

⁷Shenzhen Institute of Advanced Technology, Chinese Academy of Science, Shenzhen 518055, China.

内容摘要

本文提出了 MatterGen, 一种基于扩散模型的生成模型, 专为无机材料设计而优化。MatterGen 可以生成稳定、多样化的无机晶体材料, 并能够通过微调生成符合特定化学、对称性以及机械、电学、磁学等属性约

束的材料。相比现有的生成模型，MatterGen 生成的材料在稳定性、独特性和新颖性上均有显著提升，且生成的材料更接近能量局部最小值。研究还通过实验合成了模型生成的材料，验证了其与设计目标的接近性，展示了 MatterGen 在材料逆向设计领域的广泛应用潜力。



主要亮点

- 提出了定制化扩散过程：MatterGen 通过联合扩散处理原子类型、坐标和晶格的方式，针对晶体材料的周期性和对称性特性进行了优化，显著提高了生成材料的稳定性和多样性。
- 支持多种属性约束的逆向设计：引入适配器模块，可通过微调生成符合目标化学成分、对称性以及标量属性（如磁密度、带隙、体模量）的材料，展现了强大的条件生成能力。
- 显著优于现有生成模型：在生成稳定、独特、新颖 (S.U.N.) 材料的成功率方面，MatterGen 的表现超过了当前最先进的基线模型（如 CDVAE 和 DiffCSP），生成的材料与 DFT 能量局部最小值的偏差降低了一个数量级。
- 高效探索复杂化学体系：在三元、四元和五元化学体系中，MatterGen 生成了更多的 S.U.N.材料，尤其在复杂体系中展现出显著的效率优势，与替换法或随机结构搜索相比，所需样本量大幅减少。
- 实验验证生成材料的可合成性：合成了模型生成的 TaCr₂ O₆ 材料，实验测得的属性与设计目标相符，进一步验证了 MatterGen 生成材料的实际可行性和应用潜力。

相关链接

Paper: <https://www.nature.com/articles/s41586-025-08628-5>

Github: <https://github.com/microsoft/mattergen>

● 论文二十六

LLM360 K2: Building a 65B 360-Open-Source Large Language Model from Scratch

日期

2025-1-17

LLM360 K2: Building a 65B 360-Open-Source Large Language Model from Scratch

Zhengzhong Liu^{12*}, Bowen Tan^{3*}, Hongyi Wang^{37*}, Willie Neiswanger⁴, Tianhua Tao⁵, Haonan Li¹, Fajri Koto¹, Yuqi Wang², Suqi Sun², Omkar Pangarkar², Richard Fan², Yi Gu⁶, Victor Miller², Liqun Ma¹, Liping Tang¹, Nikhil Ranjan¹, Yonghao Zhuang³, Guowei He¹, Renxi Wang¹, Mingkai Deng³, Robin Algayres¹, Yuanzhi Li¹, Zhiqiang Shen¹, Preslav Nakov¹, Eric Xing^{13*}

¹MBZUAI ²Petuum, Inc. ³Carnegie Mellon University ⁴University of Southern California
⁵University of Illinois Urbana-Champaign ⁶University of California San Diego ⁷Rutgers University
team@llm360.ai

内容摘要

本文详细介绍了 LLM360 K2 项目，重点展示了 K2 Diamond，一款具有 65 亿参数的完全开源大型语言模型（LLM）。K2 Diamond 采用全新的 360°开源方法，覆盖从数据预处理到模型训练、微调与评估的完整生命周期，旨在推动 LLM 研究的透明性、可复现性和可访问性。K2 Diamond 通过 1.4 万亿个标记的数据进行预训练，其性能在多项基准测试中超越 LLaMA-65B 并与 Llama2-70B 相媲美，同时减少了计算成本和训练数据量。论文还探讨了训练过程中的关键技术挑战，如损失峰值、数据混合优化及安全微调，展示了模型在数学推理、代码生成和安全性对齐等领域的显著改进。

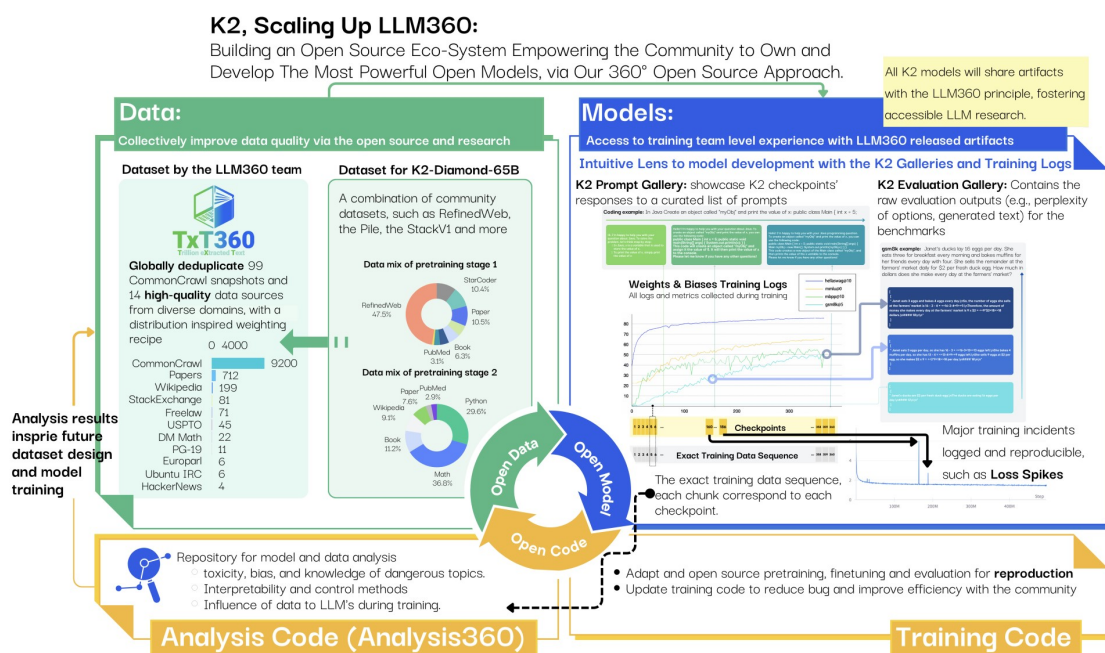


Figure 1: K2 Project Scales up the LLM360 Principles with Richer Artifacts

主要亮点

- 全面开源的 LLM 开发框架：LLM360 K2 项目秉承 360°开源原则，提供了从数据准备、模型训练到评估的全套资源，包括代码、数据集、训练日志和中间检查点，显著降低了参与门槛，促进了研究的透明性与复现性。
- 高效的数据处理与混合策略：项目采用了 TxT360 数据处理流程，整合了多种高质量数据源（如 RefinedWeb、DM-Math、StarCoder 等），并通过重复与截断策略优化数据混合比例，提高了模型的多样性与性能。
- 强大的数学推理与代码生成能力：K2 Diamond 在数学推理（如 GSM8K）和代码生成（如 HumanEval 和 MBPP）基准上表现优异，得益于在数学与代码数据上的特殊优化，如引入长上下文阶段和链式推理训练策略。
- 创新的多维并行训练方法：通过结合张量并行、流水线并行、序列并行和数据并行的 4D 并行训练策略，K2 Diamond 在 480 块 A100 GPU 上高效完成了训练，显著优化了内存利用率与计算效率。

- 改进的安全性与对齐技术: 通过构建多样化的安全微调数据集 (如直接攻击、对抗性攻击与过度拒绝提示), K2 Diamond 显著提升了对有害输入的鲁棒性, 以及对敏感话题的合理响应能力, 使其安全性达到了更高的标准。

相关链接

Paper: <https://arxiv.org/pdf/2501.07124>

Github: <https://github.com/llm360>

Model: <https://huggingface.co/LLM360/K2>

● 论文二十七

Mobile-Agent-E: Self-Evolving Mobile Assistant for Complex Tasks

日期

2025-1-20

Mobile-Agent-E: Self-Evolving Mobile Assistant for Complex Tasks

Zhenhailong Wang^{*1} Haiyang Xu^{*2} Junyang Wang² Xi Zhang²
Ming Yan² Ji Zhang² Fei Huang² Heng Ji^{*1}

¹University of Illinois Urbana-Champaign ²Alibaba Group.
^{*}Corresponding authors: Zhenhailong Wang <wangz3@illinois.edu>, Haiyang Xu <shuofeng.xhy@alibaba-inc.com>, Heng Ji <hengji@illinois.edu>.

内容摘要

本文提出了 Mobile-Agent-E, 一种能够自我进化的分层多代理移动助手, 旨在解决复杂的长步骤任务。

Mobile-Agent-E 通过显式分离高层规划和低层操作, 结合包括 Manager (任务规划)、Perceptor (视觉感知)、Operator (动作决策)、Action Reflector (动作反思) 和 Notetaker (信息聚合) 在内的多代理协作, 显著提升了任务规划和错误恢复能力。此外, 该框架引入了一个自我进化模块, 利用从任务中积累的长期记忆

(包括 Tips 和 Shortcuts) 不断优化任务执行效率和性能。与此框架配套的 Mobile-Eval-E 基准测试专注于真实世界中复杂的多应用任务，实验表明，Mobile-Agent-E 在多个基准测试上显著优于现有的最先进方法，并通过自我进化进一步提升了性能和效率。



Figure 1. We propose Mobile-Agent-E, a novel hierarchical multi-agent mobile assistant that outperforms previous state-of-the-art approaches (Zhang et al., 2023; Wang et al., 2024b;a) on complex real-world tasks. Mobile-Agent-E disentangles high-level planning and low-level action decision with dedicated agents. Equipped with a newly introduced self-evolution module that learns general Tips and reusable Shortcuts from past experiences, Mobile-Agent-E demonstrates further improvements in both performance and efficiency.

主要亮点

- 提出 Mobile-Agent-E 框架：引入分层多代理架构，显式区分高层规划与低层操作，提升了复杂任务的长远规划能力和低层动作决策的准确性。
- 引入自我进化模块：通过长期记忆中的 Tips（通用指导）和 Shortcuts（可重用操作序列）不断积累经验，改进任务执行效率和错误恢复能力，模拟人类的学习和优化行为。
- 开发 Mobile-Eval-E 基准：提出了一个新基准测试，针对复杂、长步骤、多应用交互任务，显著提高了任务难度，填补了现有研究中对复杂场景任务评估的空白。
- 显著提升性能与效率：在 Mobile-Eval-E 基准上，Mobile-Agent-E 在满意度得分上提升了 22.1%，结合自我进化后进一步提升至 33.9%，并显著降低了错误终止率，展现了卓越的规划和执行能力。

- 多模型融合与开源：支持多种大型多模态模型（LMM）作为推理代理的基础模型，并开源代码与数据，为研究社区提供了宝贵资源以进一步推动移动助手领域的发展。

相关链接

Paper: <https://arxiv.org/pdf/2501.11733>

Project: <https://x-plugin.github.io/MobileAgent/>

Code: <https://github.com/X-PLUG/MobileAgent>

Dataset: https://huggingface.co/datasets/mikewang/mobile_eval_e

● 论文二十八

DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning

日期

2025-1-20

DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning

DeepSeek-AI

research@deepseek.com

内容摘要

本文提出了首代推理模型 DeepSeek-R1-Zero 和 DeepSeek-R1，探索通过纯强化学习（RL）提升大型语言模型（LLMs）的推理能力。DeepSeek-R1-Zero 通过大规模 RL 训练，不依赖监督微调（SFT），展现了强大的推理能力，但存在可读性和语言混杂等问题。为此，DeepSeek-R1 引入了冷启动数据和多阶段训练管道，进一步提升模型性能，与 OpenAI-o1-1217 在推理任务上表现相当。此外，本文通过蒸馏技术将 DeepSeek-R1 的推理能力迁移至小型密集模型，实现了显著的性能提升并开源多个模型，推动了推理模型在研究和工业领域的广泛应用。

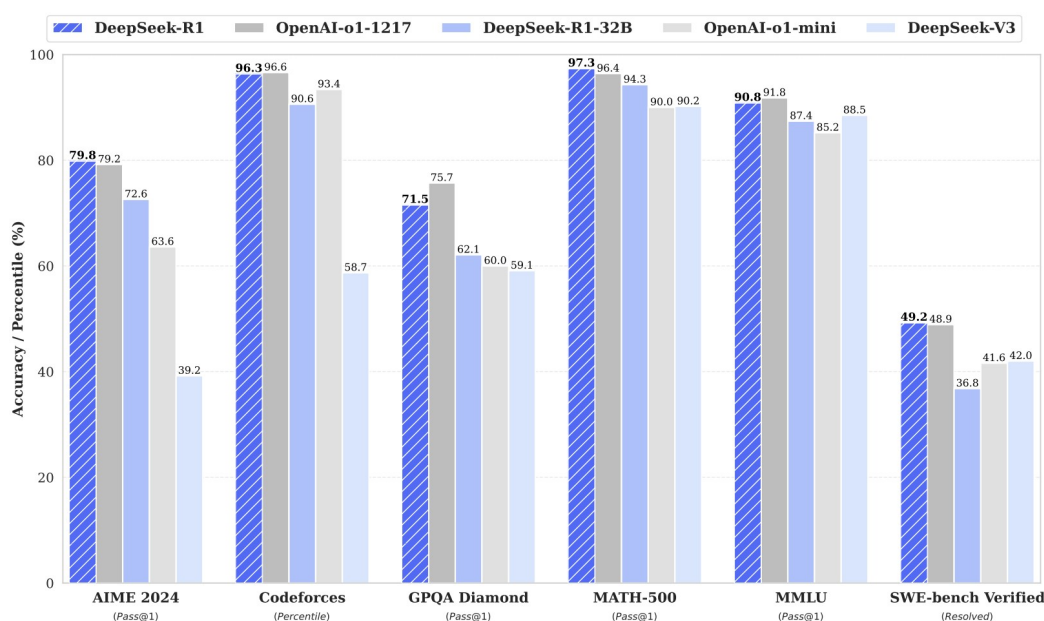


Figure 1 | Benchmark performance of DeepSeek-R1.

主要亮点

- 提出纯 RL 驱动的推理模型 DeepSeek-R1-Zero：通过 RL 直接提高模型的推理能力，无需依赖 SFT，验证 LLMs 可通过强化学习自主进化推理能力，显著提升在数学、逻辑推理等任务中的表现。
- 开发多阶段训练管道的 DeepSeek-R1：引入冷启动数据和多阶段训练方法，结合强化学习与监督微调，改善模型的可读性和多样性，进一步优化推理性能，在多个基准测试中达到或超越 OpenAI-o1-1217 水平。
- 创新的蒸馏技术：将 DeepSeek-R1 的推理能力成功蒸馏至小型密集模型（如 1.5B、7B、14B 模型），显著提升小模型的推理性能，14B 模型在多个基准测试中超越 QwQ-32B-Preview，展示了蒸馏技术的高效性和实用性。
- 广泛的基准评估与开源贡献：在多个推理与知识任务（如 AIME 2024、MATH-500、GPQA Diamond、Codeforces 等）中表现优异，并公开多种规模的模型检查点，为研究社区提供了丰富的资源和工具。
- 探索强化学习与推理能力的边界：对比蒸馏与 RL 方法的优劣，指出蒸馏技术在小模型推理能力提升中的高性价比，同时强调未来需结合更强大的基础模型与大规模 RL 探索推理能力的上限。

相关链接

Paper: https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf

Github: <https://github.com/deepseek-ai/DeepSeek-R1>

TECHNICAL UPDATES

技术动态

● 技术动态一

上海人工智能实验室进一步开源 InternVL2.5 后训练算法及训练数据

日期

2024-12-30

< Papers  arxiv:2411.10442 

Enhancing the Reasoning Ability of Multimodal Large Language Models via Mixed Preference Optimization

Published on Nov 16, 2024 ·  Submitted by  Weiyun1025 on Nov 22, 2024  #1 Paper of the day

Authors:  Weiyun Wang,  Zhe Chen,  Wenhai Wang, Yue Cao, Yangzhou Liu, Zhangwei Gao, Jinguo Zhu,  Xizhou Zhu,  Lewei Lu,  Yu Qiao,  Jifeng Dai

Abstract

Existing open-source multimodal large language models (MLLMs) generally follow a training process involving pre-training and supervised fine-tuning. However, these models suffer from distribution shifts, which limit their multimodal reasoning, particularly in the Chain-of-Thought (CoT) performance. To address this, we introduce a preference optimization (PO) process to enhance the multimodal reasoning capabilities of MLLMs. Specifically, (1) on the data side, we design an automated preference data construction pipeline to create MMPR, a high-quality, large-scale multimodal reasoning preference dataset. and (2) on the model side, we explore integrating PO with MLLMs, developing a simple yet effective method, termed Mixed Preference Optimization (MPO), which boosts multimodal CoT performance. Our approach demonstrates improved performance across multiple benchmarks, particularly in multimodal reasoning tasks. Notably, our model, InternVL2-8B-MPO, achieves an accuracy of 67.0 on MathVista, outperforming InternVL2-8B by 8.7 points and achieving performance comparable to the 10x larger InternVL2-76B. We hope this study could inspire further advancements in MLLMs. Code, data, and model shall be publicly released.

 View arXiv page

 View PDF

 Add to collection

内容摘要

基于偏好对齐的后训练算法在自然语言处理领域已引起广泛关注，且通过偏好对齐微调得到的语言模型在有用性和无害性方面表现更为优越。然而，在多模态领域，现有主流的开源模型仍依赖于预训练与有监督微调

的训练框架。由于缺乏多模态偏好对齐数据和实践经验，开源社区在偏好对齐方面的研究尚显薄弱。为解决这一问题，书生（Intern）大模型的研究团队在数据层面，构建了首个大规模多模态偏好对齐数据集 MMPR；在模型训练层面，团队提出了混合偏好优化算法（Mixed Preference Optimization, MPO），该算法旨在使模型在生成符合人类偏好的响应的同时，能够有效区分每条响应的绝对质量以及不同响应之间的相对质量。实验结果表明，基于该算法的 InternVL2.5-MPO-72B 模型在数学推理基准 MathVista 上提升了 5.9% 的性能，在文字识别基准 OCRBench 上提升了 54 分。

相关链接

技术报告: <https://huggingface.co/papers/2411.10442>

开源代码: <https://github.com/OpenGVLab/InternVL>

模型地址: https://huggingface.co/OpenGVLab/InternVL2_5-78B-MPO

数据地址: <https://huggingface.co/datasets/OpenGVLab/MMPR-v1.1>

● 技术动态二

Sam Altman 新年博文: Reflections, 预见 AGI 时代已经到来, 相信超级智能 (superintelligence) 工具可以大大加速科学发现和创新

日期

2025-1-06

Sam Altman

[« Back to blog](#)

Reflections

内容摘要

Sam Altman 在这篇博文中反思了 OpenAI 的历程及其发展。他回顾了自 OpenAI 成立以来的九年，强调了他们最初的雄心和对 AGI（人工通用智能）的信念。随着 ChatGPT 的推出，OpenAI 进入了一个快速增长的阶段，尽管这一过程充满了挑战和不确定性，Altman 表达了对团队和合作伙伴的感激之情。Altman 还谈到，在 OpenAI 的发展过程中，他亲身经历了领导危机，并从中吸取了宝贵的教训。他对未来的展望充满信心，认为在不久的将来，AI 的进步将带来超级智能工具，进而推动科学发现和社会繁荣。

相关链接:

博客: <https://blog.samaltman.com/reflections>

● 技术动态三

《MIT Technology Review》：What's next for AI in 2025

日期

2025-1-06

ARTIFICIAL INTELLIGENCE

What's next for AI in 2025

You already know that agents and small language models are the next big things. Here are five other hot trends you should watch out for this year.

By James O'Donnell, Will Douglas Heaven, & Melissa Heikkilä

January 8, 2025

内容摘要

《麻省理工科技评论》对人工智能（AI）技术在 2025 年发展热点做了如下预测：

- 1. 代理人 (Agents) 的崛起：**AI 代理人将从基本的聊天机器人发展为能够自主执行复杂任务的系统，如日程安排和软件开发。然而，这些能力可能带来重大后果，特别是在处理敏感信息时，错误可能导致严重后果。
- 2. 政府对 AI 的关注：**各国政府将更加关注 AI 的国家安全影响，例如美国限制中国获取 AI 技术。监管也将变得至关重要，欧盟的《AI 法案》可能成为全球标准。在美国，各州可能会独立采取监管行动。
- 3. AI 技术的投资与应用：**对 AI 技术的投资需要开始展示对消费者的价值，特别是在医疗保健领域，预计 AI 工具将用于诊断和药物效果监测。然而，对于在监管较少的市场（特别是全球南方）强加有缺陷的 AI 模型，存在担忧。
- 4. AI 视频技术的主流化：**AI 视频技术预计将成为主流，视频生成工具将变得更加易于获取，并通过智能眼镜技术的进步，集成到各种任务中。

在 2025 年，AI 将在技术创新、商业应用和监管环境中继续深化。AI 代理人的发展、政府监管的加强、AI 技术在医疗等领域的深入应用以及视频生成技术的普及，都是推动这一转型的关键因素。随着 AI 技术不断渗透到各行各业，全球社会、经济和政治格局将经历深刻的变革。

相关链接

文章：<https://www.technologyreview.com/2025/01/08/1109188/whats-next-for-ai-in-2025/>

● 技术动态四

Sky-T1: Train your own O1 preview model within \$450

日期

2025-1-09

	Sky-T1-32B-Preview	Qwen-2.5-32B-Instruct	QwQ	o1-preview
Math500	82.4	76.2	85.4	81.4
AIME2024	43.3	16.7	50.0	40.0
LiveCodeBench-Easy	86.3	84.6	90.7	92.9
LiveCodeBench-Medium	56.8	40.8	56.3	54.9
LiveCodeBench-Hard	17.9	9.8	17.1	16.3
GPQA-Diamond	56.8	45.5	52.5	75.2

内容摘要

UC Berkeley 的研究人员推出了一款名为 Sky-T1-32B 的开源推理大型语言模型（LLM），旨在解决现有高端 AI 模型成本高昂和封闭性的问题。Sky-T1-32B 以仅 450 美元的低训练成本实现了 32 亿参数的高效能设计，结合优化的数据扩展和参数节约技术（如稀疏计算和低秩适配 LoRA），在保持性能的同时显著降低了计算资源需求。该模型专注于推理任务，通过推理预训练和微调增强了数学推理和复杂问题解决能力，在 Math500、AIME 和 Livebench 等基准测试中超越了 OpenAI 的 o1 模型。同时，Sky-T1 的开发周期仅需 19 小时，展示了快速经济地开发高性能模型的可行性。作为开源模型，Sky-T1 的架构和训练过程完全公开，旨在推动全球研究人员和开发者的协作与创新，使 AI 技术更加普惠。其在教育、科研和自动化决策等领域的广泛适用性，结合其对环境可持续性的支持，标志着其不仅是一个技术突破，更是迈向包容性 AI 未来的重要一步。

相关链接

文章：
<https://www.marktechpost.com/2025/01/13/uc-berkeley-researchers-released-sky-t1-32b-preview-an-open-source-reasoning-llm-trained-for-under-450-surpasses-openai-o1-on-benchmarks-like-math500-aime-and-livebench/>
Github: <https://github.com/NovaSky-AI/SkyThought>
模型: <https://huggingface.co/bartowski/Sky-T1-32B-Preview-GGUF>

● 技术动态五

MiniMax-01: Scaling Foundation Models with Lightning Attention

日期

2025-1-14

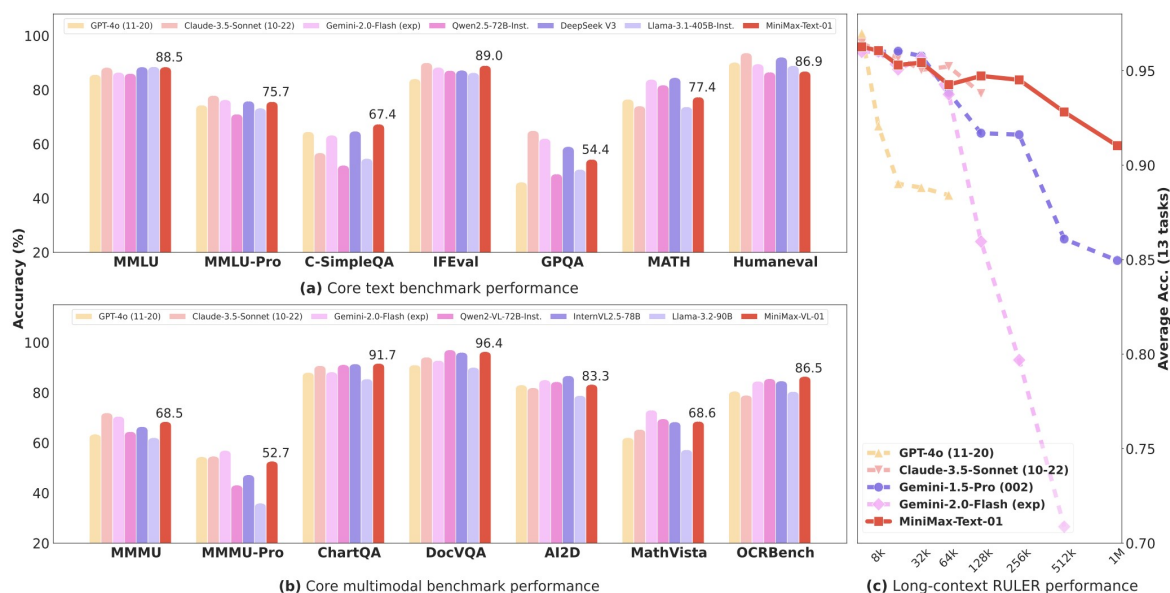


Figure 1 | **Benchmark performance.** (a) MiniMax-Text-01 on core text benchmarks. (b) MiniMax-VL-01 on core multimodal benchmarks. (c) MiniMax-Text-01 on the long-context RULER (Hsieh et al., 2024) benchmark. The performance of leading commercial and open-source models is presented for reference.

内容摘要

MiniMax-01 系列模型，包括 MiniMax-Text-01 和 MiniMax-VL-01，是一种全新的基础模型，通过创新的 Lightning Attention 和专家混合 (MoE) 技术实现了大规模扩展和高效性。这些模型以 4560 亿总参数和超长上下文窗口（训练阶段支持 100 万 tokens，推理阶段扩展至 400 万 tokens）为特色，在文本和多模态任务上表现出色。研究团队通过高效的分布式训练框架和优化的硬件利用方案，克服了传统软注意力机制的计算瓶颈，并首次成功实现了大规模线性注意力的商用部署，同时采用了混合架构将线性注意力与软注意力结合，有效提升了模型的检索和推理能力。模型在 MMLU、MathVista、LongBench 等基准测试中取得了与顶级商业模型（如 GPT-4o 和 Claude-3.5-Sonnet）相当甚至优越的性能，并在长上下文处理、推理和多模态任务上表现尤为突出。此外，MiniMax-VL-01 集成了视觉编码器，通过多阶段训练策略展现出卓越的视觉理解能力。通过开源模型权重和经济高效的 API，研究团队旨在推动基础模型的公平性和可访问性，同时明确了未来改进方向，如更高效的无软注意力架构、复杂编程任务处理能力以及更真实的长上下文评估任务，从而为构建更高效、更具扩展性和更智能的 AI 模型奠定了基础。

相关链接

文章: https://filecdn.minimax.chat/_Arxiv_MiniMax_01_Report.pdf
Github: <https://github.com/MiniMax-AI/MiniMax-01>
模型: <https://huggingface.co/MiniMaxAI/MiniMax-VL-01>
网页端: <https://hailuo.ai>

● 技术动态六

上海 AI Lab 发布 Intern3-8B-Instruct 模型

日期

2025-1-15

		同量级开源模型			闭源模型
		InternLM3-8B-Instruct	Qwen2.5-7B-Insturct	Llama3.1-8B-Instruct	GPT-4o-mini
General	CMMLU 0-shot	83.1	75.8	53.9	66.0
	MMLU 0-shot	76.6	76.8	71.8	82.7
	MMLU-Pro 0-shot	57.6	56.2	48.1	64.1
Reasoning	GPQA-Diamond 0-shot	37.4	33.3	24.2	42.9
	DROP 0-shot	83.1	80.4	81.6	85.2
	HellaSwag 10-shot	91.2	85.3	76.7	89.5
	KOR-Bench 0-shot	56.4	44.6	47.7	58.2
MATH	MATH-500 0-shot	83.0*	72.4	48.4	74.0
	AIME2024 0-shot	20.0*	16.7	6.7	13.3
Coding	LiveCodeBench (2407-2409)Pass@1	17.8	16.8	12.9	21.8
	HumanEval Pass@1	82.3	85.4	72.0	86.6
Instruction	IFEval Prompt-Strict	79.3	71.7	75.2	79.7
LongContext	RULER (4-128K) Average	87.9	81.4	88.5	90.7
Chat	AlpacaEval 2.0 LC WinRate	51.1	30.3	25.0	50.7
	WildBench Raw Score	33.1	23.3	1.5	40.3
	MT-Bench-101 Score(1-10)	8.59	8.49	8.37	8.87

* Thinking Mode

内容摘要

上海人工智能实验室发布了书生浦语 3.0 大模型(InternLM3),通过创新的数据精炼框架和思维密度(IQPT)提升,显著提高了模型性能和训练效率。该模型仅用 4T 预训练数据实现了主流开源模型 18T 数据的效果,训练成本降低 75%以上,数据思维密度较同量级模型提升 4 倍以上。书生·浦语 3.0 首次在通用模型中融合了常规对话和深度思考能力,支持一键切换模式,适应多任务场景需求。其通过智能化数据处理和高价值数据合成技术,构建了高质量指令数据集,并基于世界知识树和多智能体方法优化用户交互体验。模型在 CMMLU、GPQA 等权威评测中表现优异,综合性能接近 GPT-4o-mini。作为开源模型,书生·浦语 3.0 支持快速部署与应用,并通过与国产算力厂商合作实现高效推理,推动 AI 生态繁荣。它不仅展现了卓越的推理与创作能力,还成为开源社区首个支持浏览器操作的对话模型,进一步拓展了智能体任务的深度应用。

相关链接

Github: <https://github.com/InternLM/InternLM>
模型: <https://huggingface.co/internlm>

● 技术动态七

OpenAI 发布首个 L3 级智能体

日期

2025-1-23

Benchmark type	Benchmark	Computer use (universal interface)		Web browsing agents	Human
		OpenAI CUA	Previous SOTA	Previous SOTA	
Computer use	OSWorld	38.1%	22.0%	-	72.4%
Browser use	WebArena	58.1%	36.2%	57.1%	78.2%
	WebVoyager	87.0%	56.0%	87.0%	-

内容摘要

OpenAI 正式发布了首个 L3 级智能体 Operator,通过其背后的核心技术 Computer-Using Agent(CUA),实现了 AI 直接与图形用户界面 (GUI) 交互的能力,突破了传统依赖 API 的局限性。CUA 能够像人类一样操

作屏幕上的按钮、菜单和文本框，完成包括在线订餐、商品购买、表单填写等多步骤复杂任务，并具备纠错和适应意外变化的能力。这一技术融合了 GPT-4o 的视觉感知、深度推理以及强化学习方法，显著提升了任务成功率：在 WebArena 和 WebVoyager 的测试中分别达到了 58.1% 和 87% 的成功率，刷新了多项 SOTA 记录。然而，与人类表现仍有差距，特别是在复杂任务和更高效的操作上，CUA 的可靠性和精确性在某些场景中还需要进一步优化。此外，CUA 在执行任务时会实时进行屏幕截图与内在推理，确保任务完成的可追踪性和安全性，同时在人类接管时保证隐私。尽管目前仅限美国 Pro 用户体验，Operator 的发布被认为是通向 AGI 的重要里程碑，标志着 AI 在人机交互和数字世界操作方面迈出了革命性的一步，未来其动作空间还将进一步扩展，并计划开放 API 让开发者构建定制化智能体。与此同时，OpenAI 通过强调安全性，确保 CUA 不会执行有害或危险任务，并通过用户反馈不断改进其应用场景和可靠性。

相关链接

文章：https://cdn.openai.com/operator_system_card.pdf

● 技术动态八

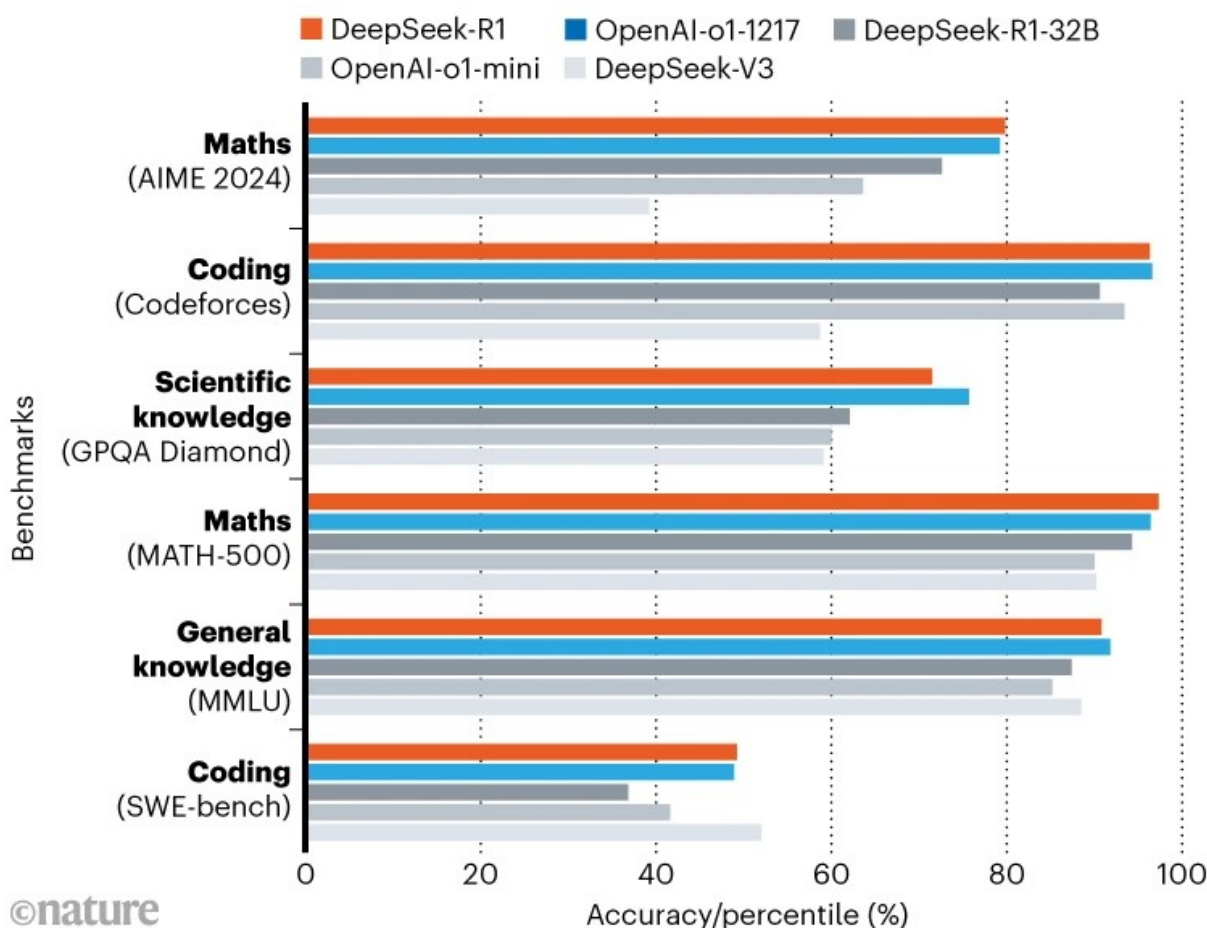
China's cheap, open AI model DeepSeek thrills scientists

日期

2025-1-23

AI RIVALS

DeepSeek tested three versions of its large language models against OpenAI's o1 models on mathematics, coding and reasoning tasks. DeepSeek-R1 beat or rivalled o1 on maths and coding benchmarks.



内容摘要

中国公司 DeepSeek 推出了新一代大型语言模型 DeepSeek-R1，作为一种开放且高性价比的“推理”模型，引起了科学界的广泛关注。DeepSeek-R1 通过“思维链”方法提升了复杂任务的解决能力，在数学、化学和编程等多个领域的性能与 OpenAI 的 o1 模型相当，同时以显著更低的成本运行，且其开源特性使研究人员能够深入研究和改进模型。DeepSeek 利用强化学习对其上一代 V3 模型进行微调，通过奖励正确答案和清晰的推理过程优化模型性能，并创新性地采用了“专家混合”架构和资源节约算法。在基准测试中，DeepSeek-R1 在 MATH-500 数学任务中取得了 97.3% 的高分，并在 Codeforces 编程竞赛中超越 96.3% 的人类选手，与 o1 表现相当。尽管其部分能力略低于 o1，但在某些领域如量子光学计算上表现出色。由于采用 MIT 许可，研究人员可以复用和改进该模型，但其训练数据尚未完全公开。分析认为，DeepSeek 在有限计算资源条件下的成

功体现了算法创新的重要性，削弱了美中在 AI 领域的差距，同时其开放性也为模型推理过程的可解释性提供了新的可能性。DeepSeek-R1 的发布展示了中国在人工智能领域的创新潜力，并为科学研究和 AI 开发提供了一种高效且透明的工具。

相关链接

文章: <https://www.nature.com/articles/d41586-025-00229-6>

RESOURCE SHARING

资源分享

● 资源分享一

Foundations of Large Language Models

类型：电子书

内容提要

《Foundations of Large Language Models》是由东北大学 NLP 实验室撰写的一本 231 页的书，旨在系统介绍大型语言模型（LLMs）的核心概念与基础技术。全书分为四章，涵盖预训练、生成式模型、提示方法和对齐技术等内容。书中首先回顾了预训练的背景与方法，阐述了基于 Transformer 架构的自监督学习如何推动通用语言理解与生成的进步，并以 BERT 和 GPT 为例展示了预训练模型的强大能力。随后，书中探讨了生成式模型的关键步骤及其在大规模训练和长文本处理中的扩展能力，强调了 LLMs 在自然语言处理中的广泛应用与潜力。提示方法部分聚焦于通过设计提示来提升模型性能，探讨了上下文学习和自动提示工程等技术，展示了 LLMs 在无需微调的情况下适应新任务的能力。对齐部分则深入分析了如何通过指令对齐和人类偏好对齐，使模型输出更符合用户需求，并解决潜在的伦理和偏见问题。整本书以笔记式的风格编写，旨在为读者提供灵活的学习路径，无论是深入研究某一领域还是全面了解 LLMs，都能从中获得启发与指导。

资源链接

Paper: <https://arxiv.org/pdf/2501.09223>

人工智能前沿快报

Artificial Intelligence
Newslettler



2025 年第 01 期

总第 19 期

广东省图象图形学会

主办

本期编委

广东省图象图形学会

金连文	华南理工大学
谭明奎	华南理工大学
钟亦武	香港中文大学
林上港	华南农业大学
李斌	深圳大学
谢晓华	中山大学
王泽宇	香港科技大学 (广州)
李冠彬	中山大学
熊龙飞	金山办公
段纪伟	金山办公
单泽昱	华南理工大学
吴诗航	华南理工大学
许毅平	华南农业大学
陈嘉杜	华南农业大学

人工智能 前沿快报

2025 年第 01 期

总第 19 期

— 2025. 01.26

声明：

- (1) 本快报内容提要、文章亮点等部分内容由 AI 生成，不一定准确及全面，文章完整内容及思想应以原文为准。
- (2) 本文大部分插图来源于相关论文或文章，版权归原作者或原出版社所有。
- (3) 本快报摘录文章观点不代表编委会及主办方立场。



学会官方公众号

广东省图象图形学会

GDSIG

学术交流、技术咨询、科学普及