

MiniMax-M1 模型技术报告

📅 2025年7月2日 ⌚ 1 分钟阅读

#AI #MiniMax #M1

MiniMax-M1 模型技术报告

M1 模型的Tech Report总结

MiniMax-M1模型是MiniMax公司推出的首个开源、大规模混合注意力推理模型。它在技术上融合了多种先进特性，旨在高效处理长上下文并进行复杂的推理。

以下是MiniMax-M1模型的主要技术特点：

1. 混合专家 (MoE) 架构

MiniMax-M1是在其前身MiniMax-Text-01模型的基础上开发的。

它总共包含 **4560亿个参数**，其中每个token激活 **459亿个参数**。

该模型集成了32个专家（experts），以最大化计算容量。

2. 闪电注意力 (Lightning Attention) 机制

MiniMax-M1的核心在于结合了混合专家（MoE）架构和闪电注意力机制。闪电注意力是线性注意力的一种I/O感知优化实现。

其架构设计独特：每七个使用闪电注意力的TransNormer块之后，会跟随一个使用softmax注意力的Transformer块。这种混合设计旨在平衡计算效率和检索能力。

高效的测试时计算扩展：与传统的Transformer架构相比，MiniMax-M1在生成长序列时能显著降低计算成本。例如，在生成长度为10万个token时，M1的FLOPs消耗约为DeepSeek R1的25%。这使得它在推理和大规模强化学习（RL）训练中效率更高。

目录

文章信息

字数

阅读时间

发布时间

更新时间

标签

#AI #MiniMax #M1

超长上下文支持：MiniMax-M1模型原生支持高达 **100万个 token的上下文长度**，并且输出生成长度可达 **8万个token**。这比DeepSeek R1的上下文大小长8倍，比目前所有开源LRM长一个数量级。

3. 高效的强化学习 (RL) 训练框架

MiniMax-M1通过大规模强化学习进行训练，涵盖了从传统数学推理到基于沙盒的真实世界软件工程环境等多样化问题。

CISPO算法：MiniMax-M1引入了一种新颖的RL算法——**CISPO (Clipped IS-weight Policy Optimization)**。该算法通过裁剪重要性采样权重而非token更新来稳定训练，相比GRPO和DAPO等其他竞争性RL变体，在效率上表现更优。

针对混合架构的RL挑战解决方案：研究团队解决了在混合注意力架构下进行RL扩展时遇到的独特挑战，例如训练模式和推理模式之间计算精度的不匹配问题。通过将LM输出头的精度提高到FP32，成功地使训练和推理的概率重新对齐，从而实现了奖励的增长。

优化器超参数敏感性：针对梯度幅度范围广和相邻迭代梯度相关性弱的问题，MiniMax-M1调整了AdamW优化器的参数设置 ($\beta_1 = 0.9$, $\beta_2 = 0.95$, $\text{eps}=1\text{e-}15$) 以确保训练收敛性。

重复检测与早期截断：为了防止病态冗长和重复的响应导致模型不稳定，模型实现了基于token概率的启发式早期截断规则，当连续3000个token的概率超过0.99时停止生成。

高效的训练成本：结合混合注意力架构和CISPO算法，MiniMax-M1的全程RL训练仅需3周，使用512个H800 GPU，租用成本约为53.47万美元。

4. 多阶段训练策略

持续预训练：在MiniMax-Text-01的基础上，通过额外7.5万亿 token的精心策划、推理密集型语料库进行持续预训练，以增强其内在推理能力和长上下文处理能力。

监督微调 (SFT)：进行冷启动监督微调，注入链式思考 (CoT) 等特定推理模式，为后续RL阶段打下坚实基础。

分阶段上下文扩展：为了应对混合闪电架构模型在训练长度扩展时可能出现的梯度爆炸问题，采用了更平滑的分阶段上下文长度扩展策略，从32K逐渐扩展到1M token的训练上下文。

长思考能力扩展：训练了两个版本的MiniMax-M1模型，最大生成长度分别为4万和8万个token，其中8万个token的模型进一步证明了扩展测试时计算的好处。

5. 多样化的RL训练数据与奖励模型

基于规则可验证的任务：包括数学推理、逻辑推理、竞技编程和软件工程问题。其中，软件工程数据利用GitHub真实世界的issues和PRs构建，并通过容器化沙盒环境执行代码以提供可验证的反馈。

基于模型反馈的通用领域任务：包括科学、技术、工程和数学(STEM) 等客观性问题，以及指令遵循、创意写作等没有明确标准答案的任务。利用生成式奖励模型（GenRM）提供反馈，并采取措施减少长度偏见。

课程学习与动态权重：RL训练采用精心设计的课程和动态加权策略，首先侧重于基于规则可验证的推理密集型任务，然后逐步混合通用领域任务，以确保模型在专业技能和通用泛化能力之间取得平衡。

6. 性能表现

MiniMax-M1在软件工程、工具使用和长上下文任务中表现出特别的优势。

在agentic工具使用基准TAU-Bench上超越了Gemini 2.5 Pro。

在长上下文理解基准上超越了OpenAI o3和Claude 4 Opus。

在复杂场景（如软件工程、长上下文和工具使用）中表现突出，与DeepSeek-R1-0528等领先的开源模型竞争，甚至在某些方面表现更优。

MiniMax-M1及其相关资源（GitHub代码库、Hugging Face模型、vLLM和Transformers部署指南、商业API）已公开，旨在促进人工智能领域的合作与进步。

7. 行业领先的评测表现

SWE-bench

MiniMax-M1-40k/80k分别取得55.6%/56.0%成绩，略低于DeepSeek-R1-0528，但遥遥领先其它开源模型。

长上下文理解

超越OpenAI o3与Claude 4 Opus，仅次于Gemini 2.5 Pro，全球排名第二。

工具使用 (TAU-bench)

MiniMax-M1-40k领跑所有开源权重模型，并战胜Gemini 2.5 Pro。

扩展性验证

80k版本在大多数基准测试中始终优于40k，证明大规模扩展的有效性。

8. 开放生态与部署支持

权重与技术报告

官方Hugging Face与GitHub账号开放下载，支持透明复现和二次开发。

推理部署

vLLM和Transformer项目均已支持，SGLang也在推进中，便于各类生产环境落地。

9. 商业化与开放性

不限量免费体验

MiniMax APP和Web端不限量免费用，极大降低开发者门槛。

API定价超低

0-32k输入：0.8元/百万Token（输出8元/百万Token）

32k-128k输入：1.2元/百万Token（输出16元/百万Token）

128k-1M输入：2.4元/百万Token（输出24元/百万Token）

高性价比，尤其128k-1M输入段DeepSeek不支持。

分享
这篇
文章



相关文章推荐

MiniMax

MiniMax

**Andrej
Karpathy...**

Andrej Karpathy
在各个场合的...

**Ilya
Sutskever...**

Ilya Sutskever 在
各个场合的观点