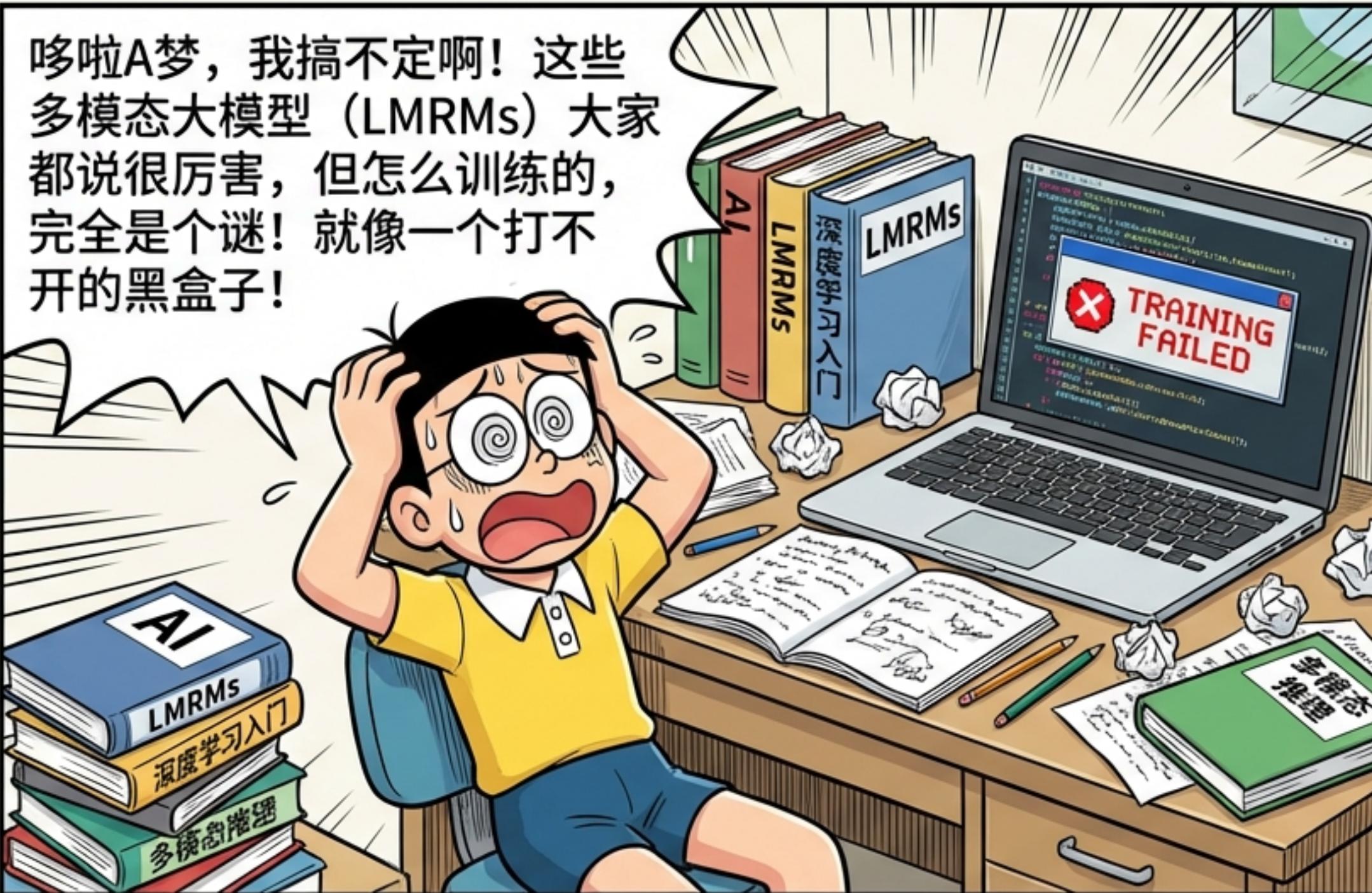


哆啦A梦的挑战：打开多模态推理的“黑盒子”



锵锵！就是这个——“开放式多模态推理秘笈”！它提供了一套完全透明、可复现的两阶段训练方法，任何人都可以跟着学！

チャチャーン！

第一阶段：监督微调（SFT）

- 打下坚实的推理基础

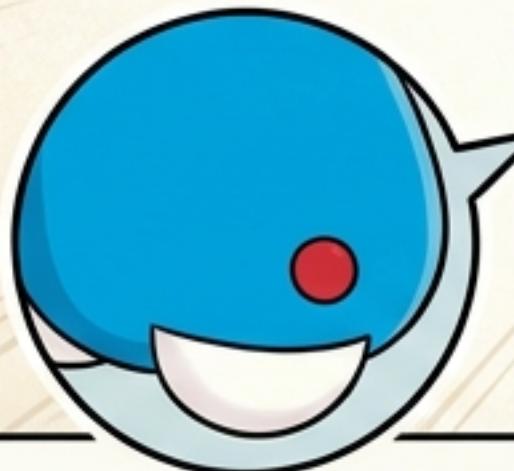
第二阶段：强化学习（RL）

- 精进和稳定推理能力

我们的秘笈
是第一个把所有流程、
数据和模型全部
公开的！

我们公开了一切！

方法	数据流程	SFT数据	RL数据	模型
Qwen2.5-VL-7B	✗	✗	✗	✓
OVR	✗	✓	✗	✓
OpenMMReasoner (我们)	✓	✓	✓	✓



第一阶段：监督微调（SFT） - 为顶尖推理能力打好基础

SFT阶段就像准备一道大餐的备料环节。我们的目标是创建一个高质量、高多样性的数据集，为模型打下坚实的推理基础。我们从10.3万个原始问题开始，最终制作出87.4万份高质量的“学习材料”！



10.3万
原始问题



数据蒸馏与扩展



跨领域混合



87.4万
最终SFT数据集

秘诀一：一位好老师至关重要

教师模型对性能的提升

基线模型

45.3%

教师模型A
(Qwen2.5-VL-72B)

+4.5
49.8%

教师模型B
(Qwen3-VL-235B)

+5.2
50.5%



Qwen3-VL-235B

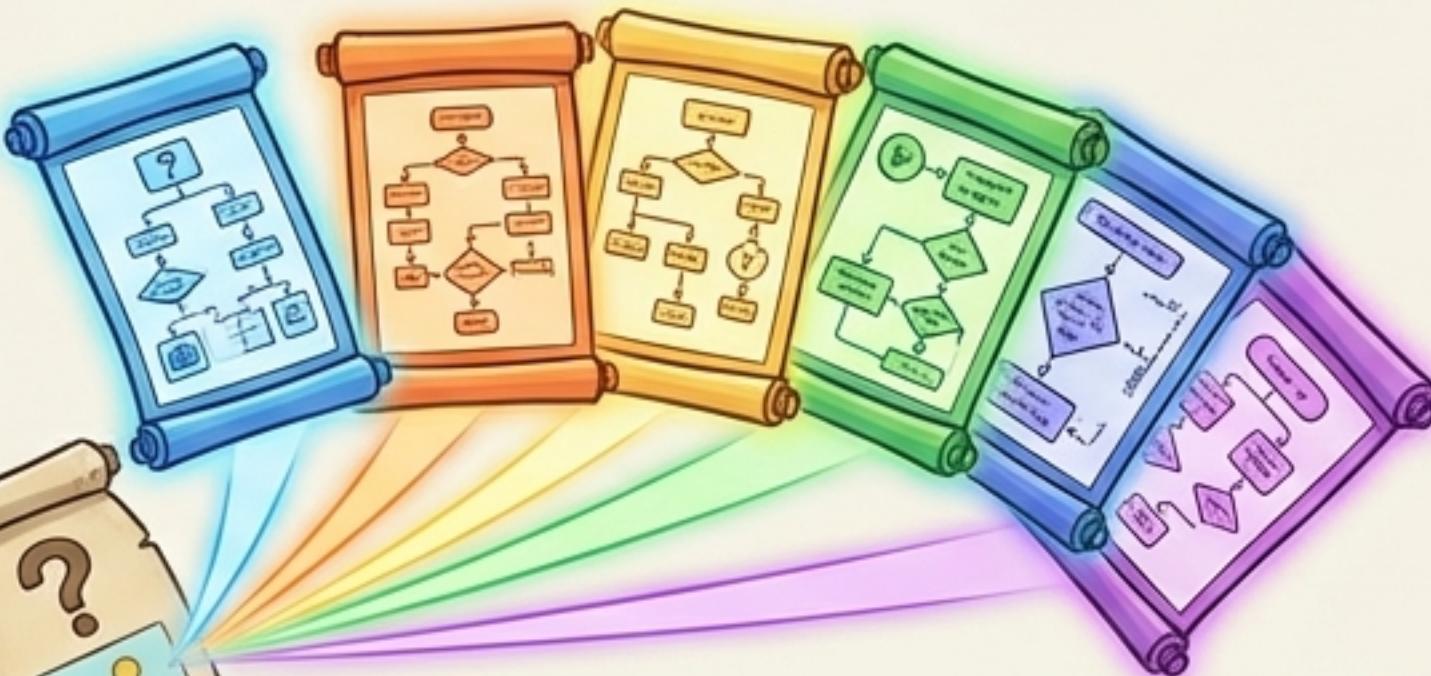
Qwen2.5-VL-72B

要制作最优质的学习材料，
我们需要一位顶尖的‘老师’来提供
答案范例。我们发现，老师越强，学生
(我们的模型) 的进步就越快！

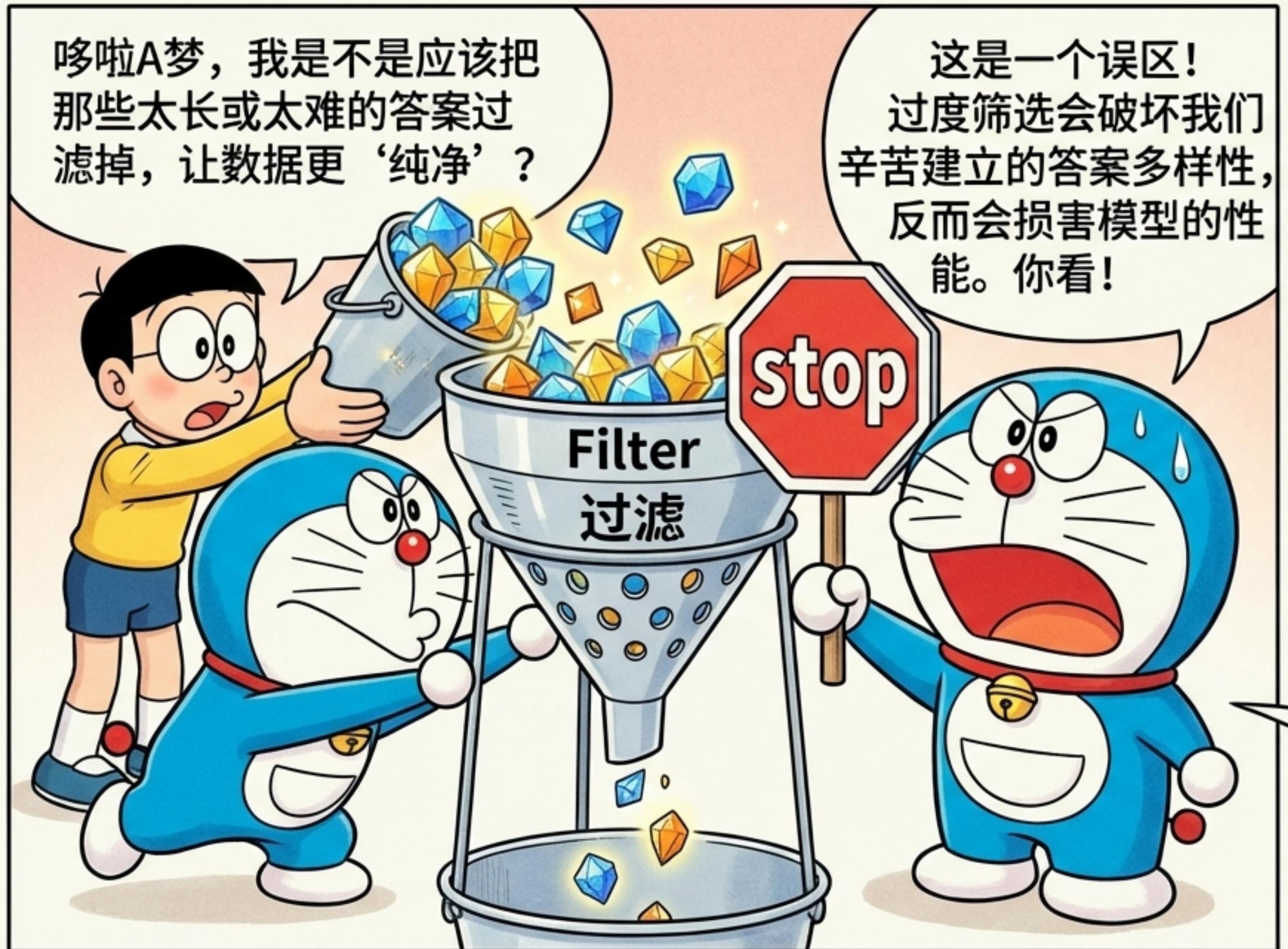
所以，我们选择了最强的
Qwen3-VL-235B作为
我们的‘特级教师’！

秘诀二：答案的多样性远比数量更重要

“记住，大雄！同一个问题，要从不同角度思考！我们让老师针对每个问题生成多个经过验证的答案，这能极大地增强模型的泛化能力。”



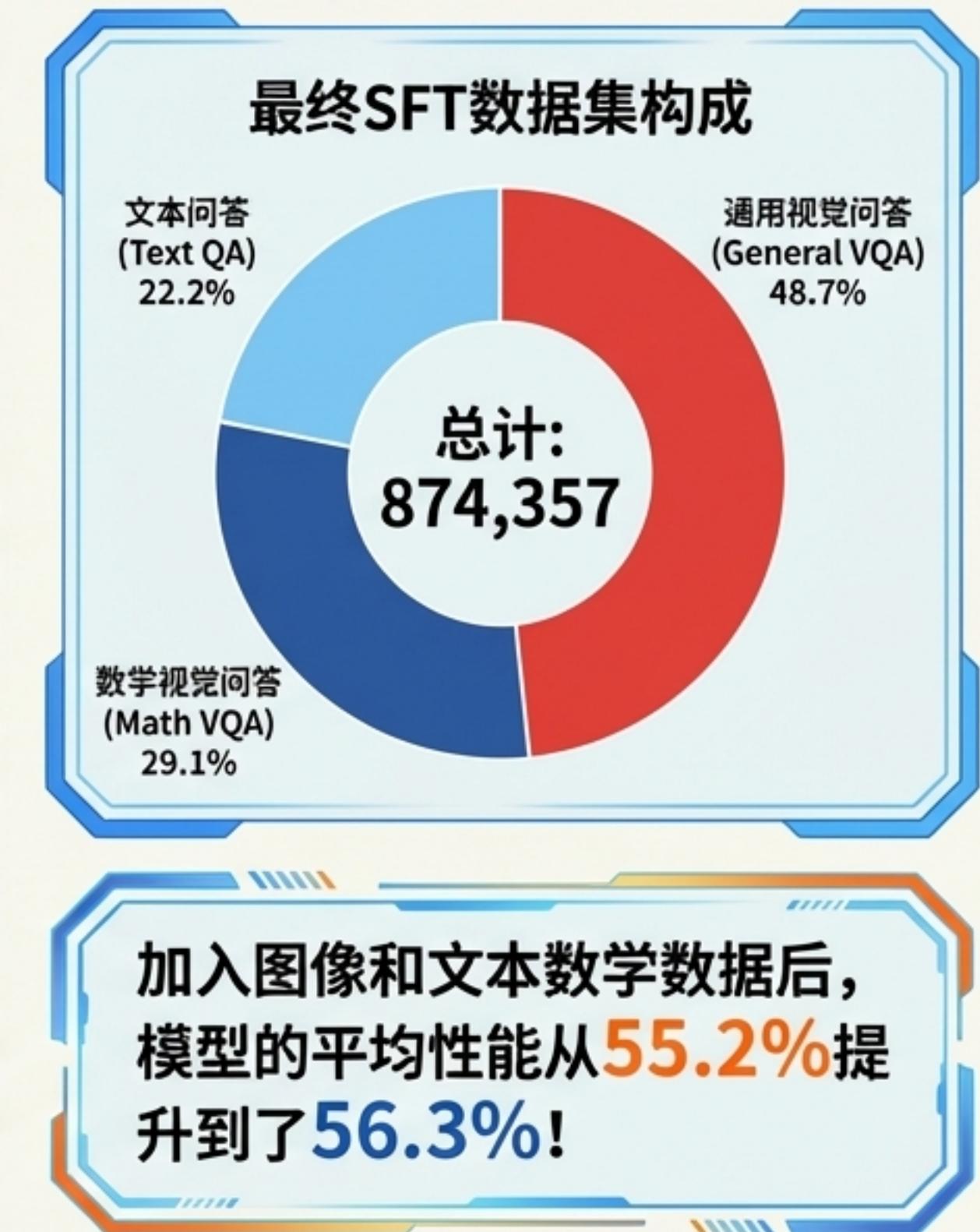
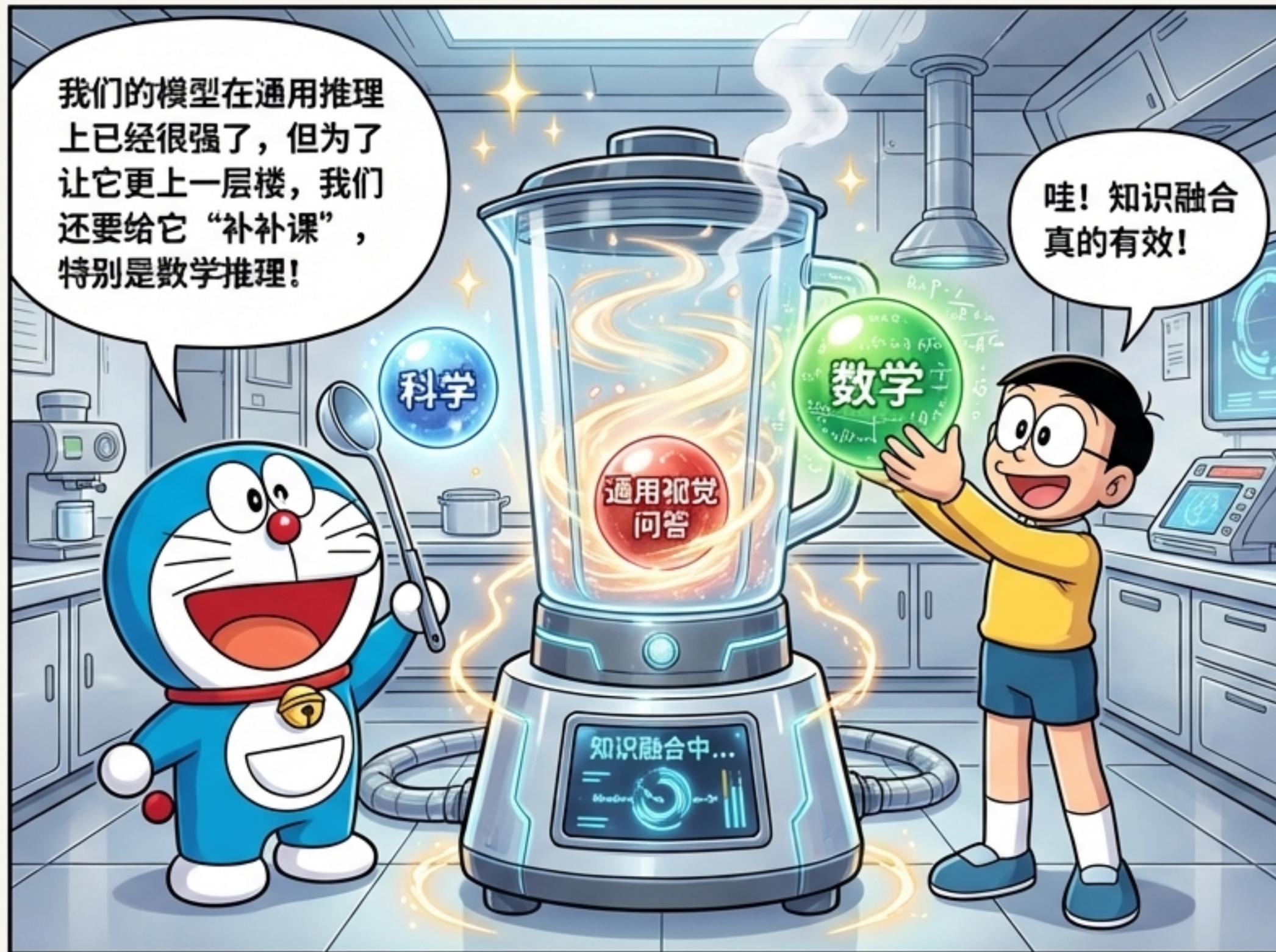
一个常见的陷阱：不要过度筛选数据！



策略	平均性能
无筛选	55.2%
长度筛选	54.2%
难度筛选	51.3%

所以，我们的最终配方是：保留所有经过验证的答案，不去画蛇添足。

最后一步：融合跨领域知识，打造全能选手



第二阶段：强化学习 (RL) - 从优秀到卓越的最后冲刺

“大雄，SFT已经为我们打下了坚实的基础。现在是高手进阶课时间！通过强化学习，我们可以进一步‘打磨’模型，让它的推理能力更精准、更稳定。”

锐化 (Sharpen)

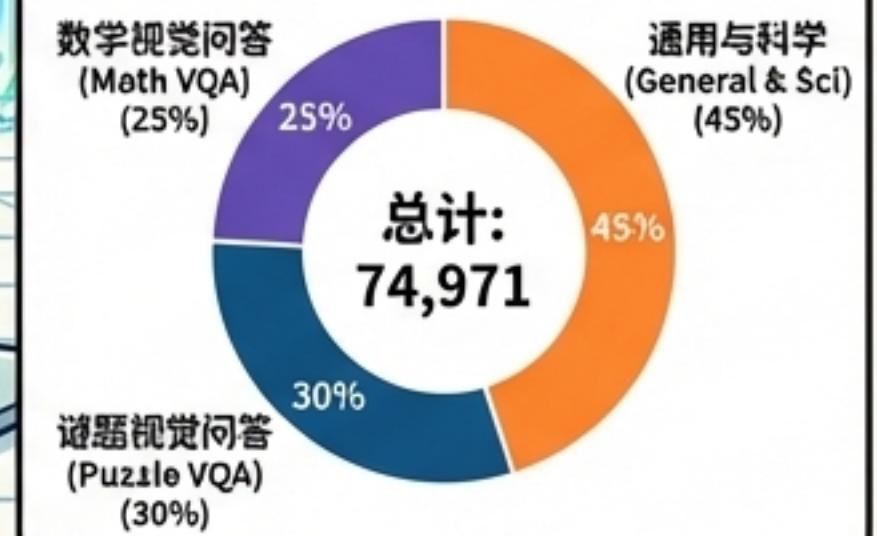
提升推理的准确性。



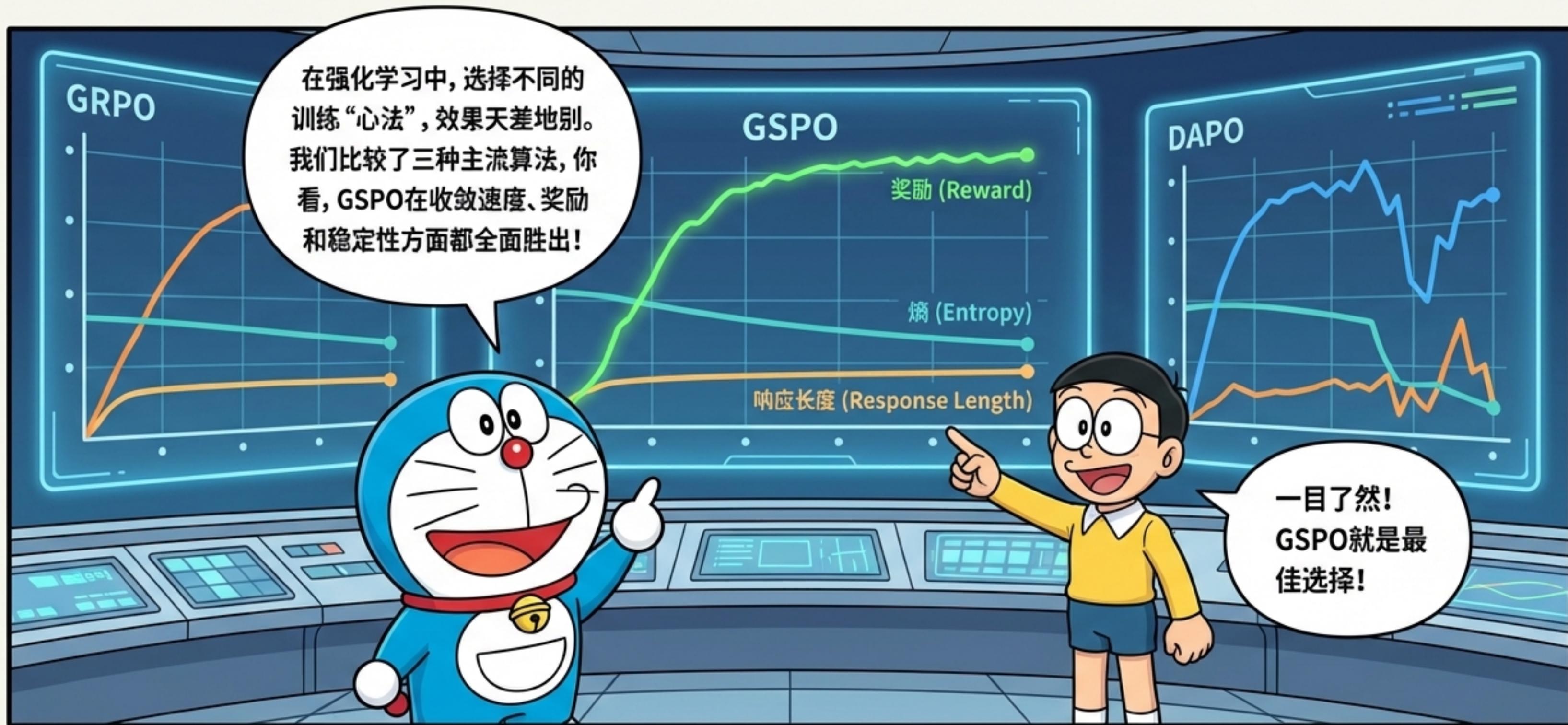
稳定 (Stabilize)

确保在各种任务下都能有可靠表现。

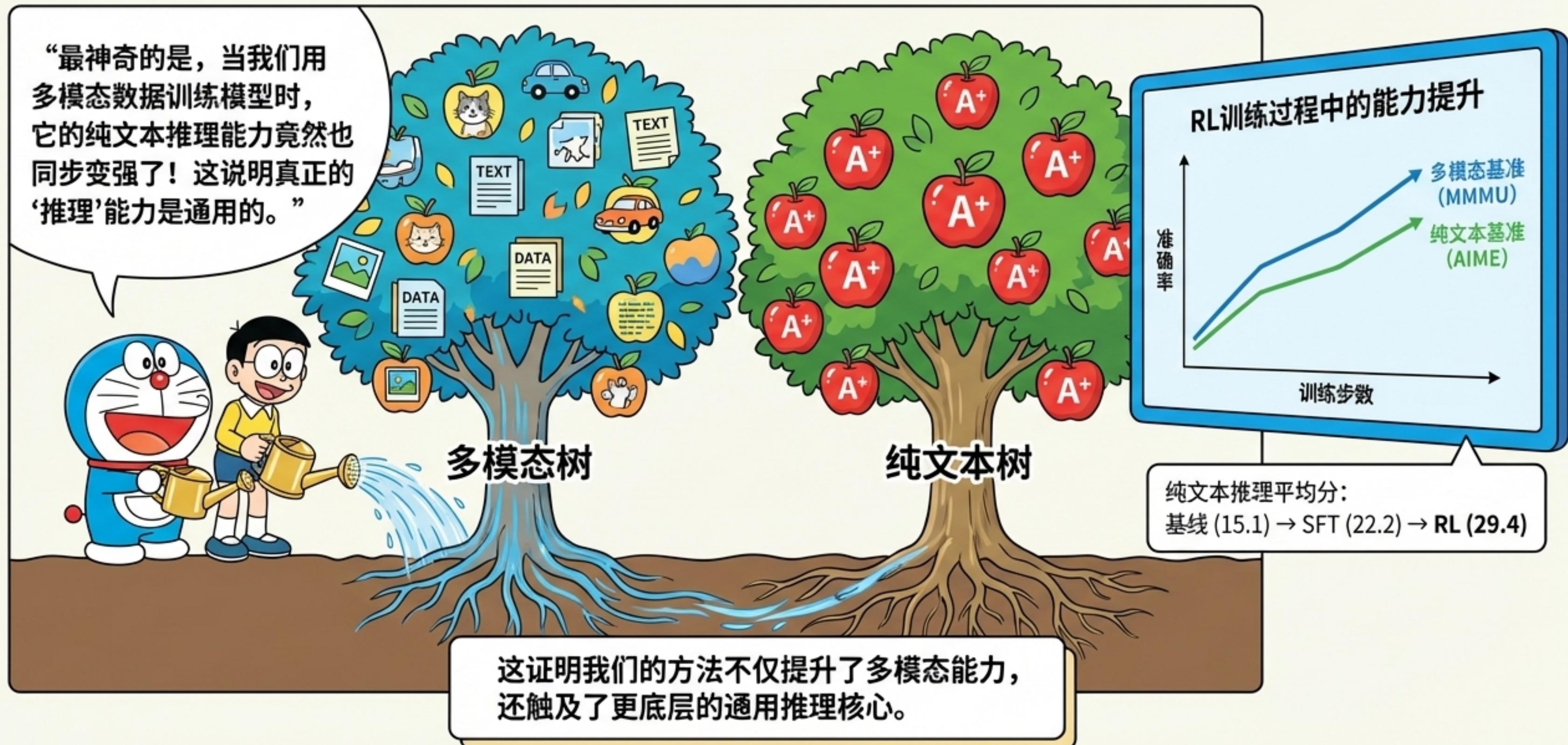
RL数据集构成



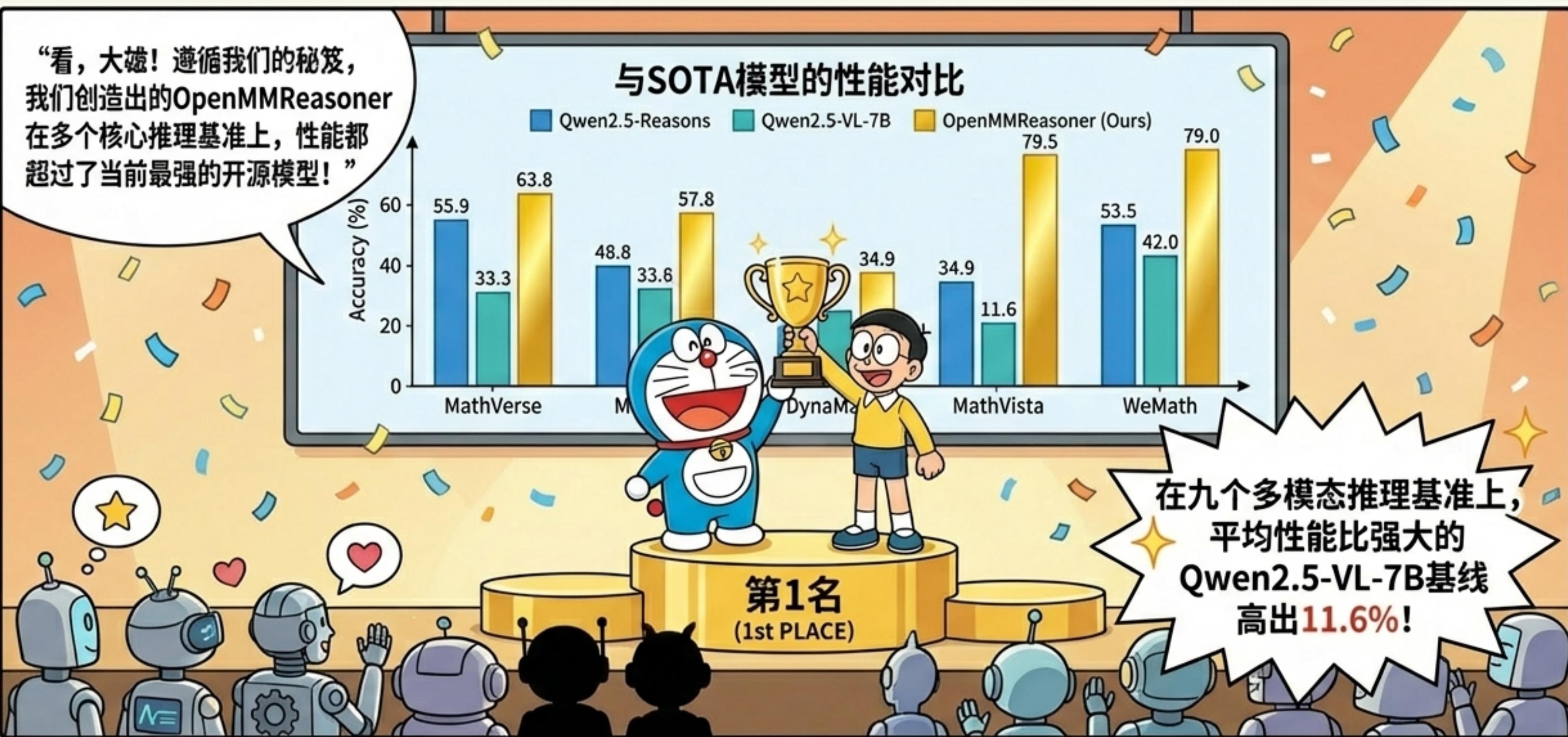
RL的成败关键：选择最高效的训练算法



意外之喜：推理能力实现了跨模态迁移！



成果揭晓：OpenMMReasoner全面超越顶尖模型



深入数据：在各大视觉推理基准上的详细表现

对于想深入了解的同学，这里是详细的成绩单。无论是在数学、逻辑还是常识推理上，我们的模型都表现出色。



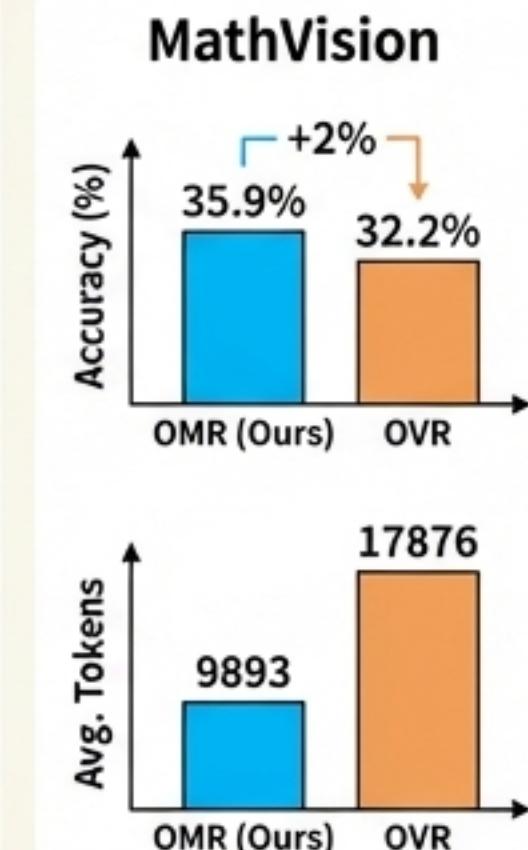
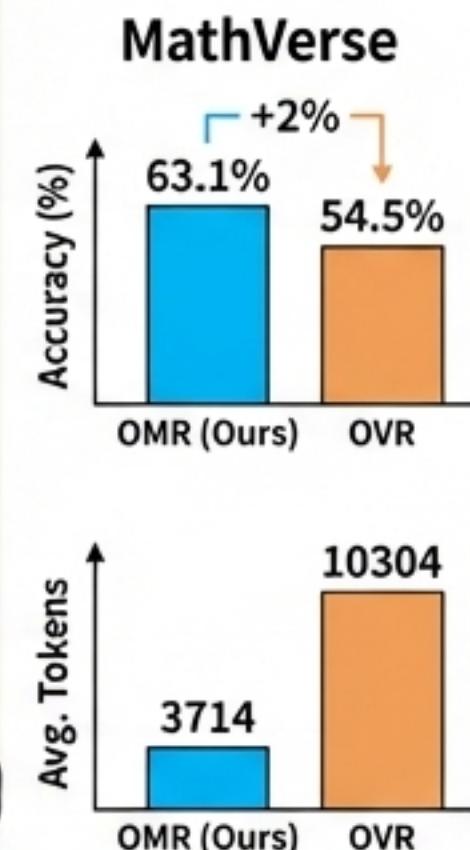
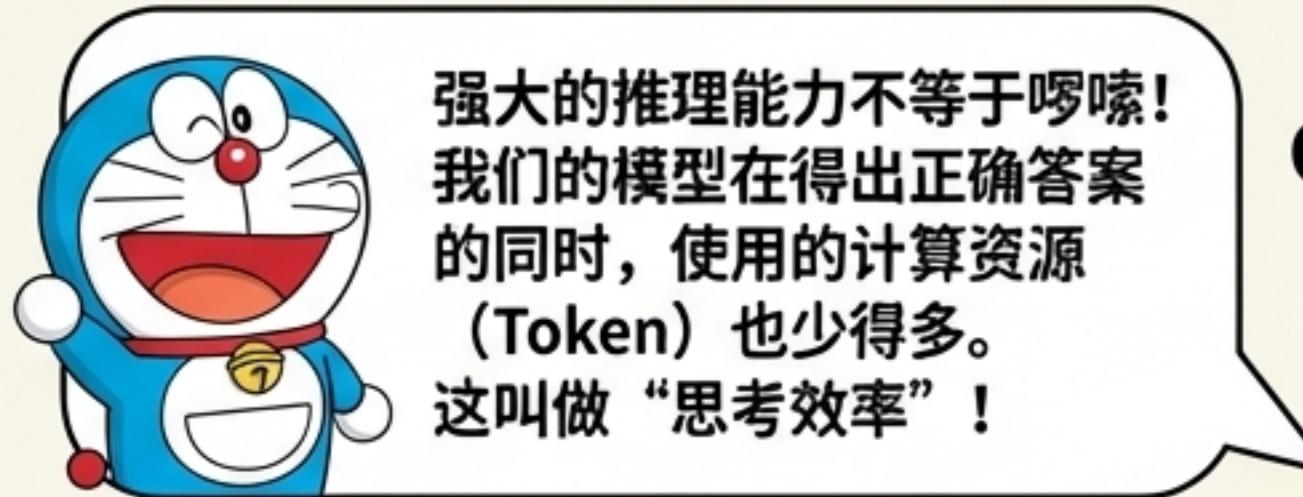
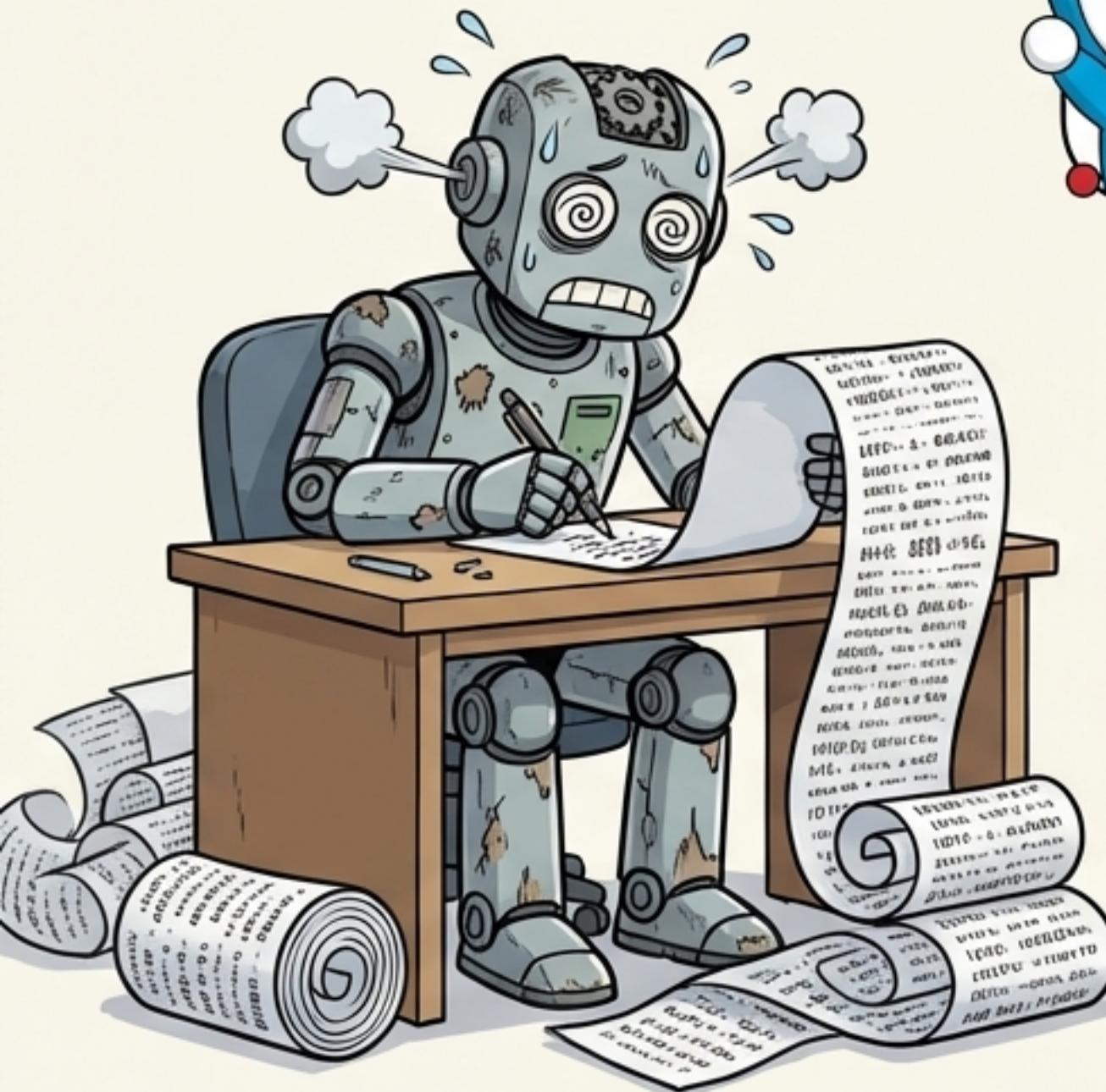
Model	MathVista	MathVerse	WeMath	MMU
SFT Methods				
LLaVA-OneVision-1.5-8B	77.1	51.4	71.0	52.6
OMR-7B-ColdStart (SFT)	79.7	63.8	77.0	57.8
RL-based Methods				
OVR-7B	79.2	63.8	79.0	54.7
OMR-7B (RL)	79.5	63.8	79.0	57.8



小结：我们的两阶段方法，无论是在SFT后还是RL后，都展现出了强大的竞争力。

不仅更强，而且更高效：实现推理深度与效率的平衡

其他模型（如 OVR）



OpenMMReasoner (我们)



用更少的思考，办更漂亮的事！

哆啦A梦的推理模型大师课总结



这份秘笈属于每一个人：开放，共建，探索未来



“太棒了！哆啦A梦，有了这份秘笈，大家都能做出超强的推理模型了！”

“没错！科学的真谛就在于分享。这就是我们把它完全开源的原因。希望它能成为大家研究的坚实基础！”

- 我们开源了所有代码、数据流程和模型权重。
- 一个完全透明、可复现的SFT+RL框架。

立刻访问我们的项目，开始你的多模态推理之旅！

<https://github.com/EvolvingLMMs-Lab/OpenMMReasoner>

