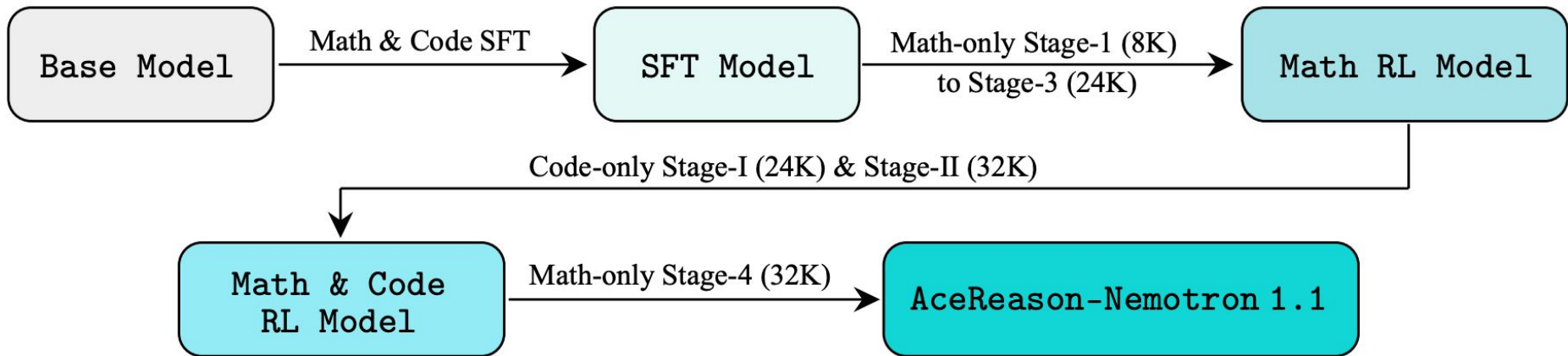




# AceReason-Nemotron 1.1: Advancing Math and Code Reasoning through SFT and RL Synergy

Zihan Liu, Zhuolin Yang, Yang Chen, Chankyu Lee, Mohammad Shoeybi, Bryan Catanzaro, Wei Ping

# Training Framework (SFT + RL)



- Cold Start: Math & Code SFT
- RL: separate math and code training
- RL: a stage-wise response length extension
  - 8K -> 16K -> 24K -> 32K

# SFT Research Questions

- How does scaling SFT data size improve model accuracy on math and code benchmarks?
  - Is it more effective to scale the number of unique prompts or the number of responses per prompt? What is the best strategy in practice?

# Scaling SFT Dataset Size

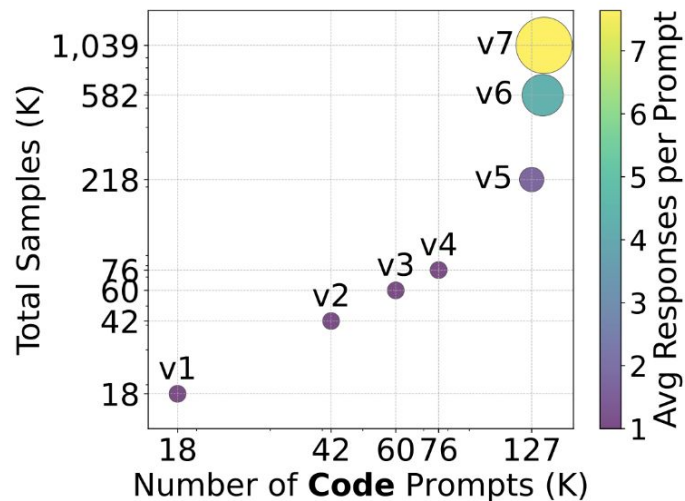
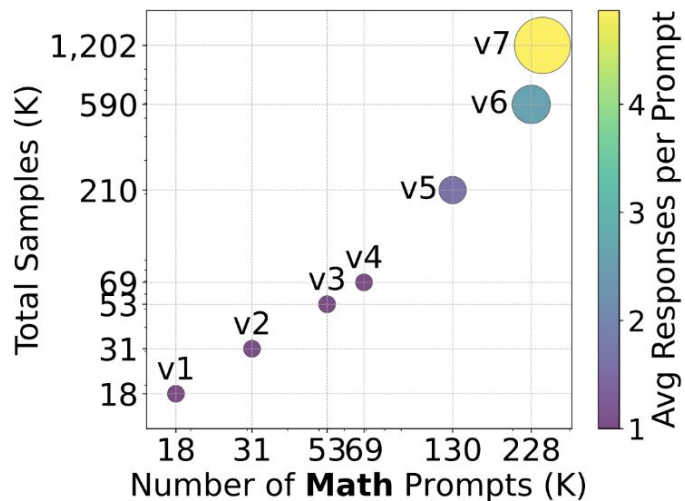


Figure 4: Log-scaled data statistics for the number of math and code prompts and the average number of responses per prompt. Each SFT dataset consist of both math and code SFT samples.

SFT sample responses are all generated by DeepSeek-R1

# SFT Research Questions

- Is it more effective to scale the number of unique prompts or the number of responses per prompt? What is the best strategy in practice?
  - Using least square methods to run the estimation
  - Scaling the number of unique prompts has higher impact

# RL Analyses

- RL from different SFT models
- When to do overlong filtering in RL
  - Whether to ignore samples having responses longer than the context length
- How training temperature affects the progress of RL
- Importance of Stage-1 (8K)
- Math-RL improves coding performance

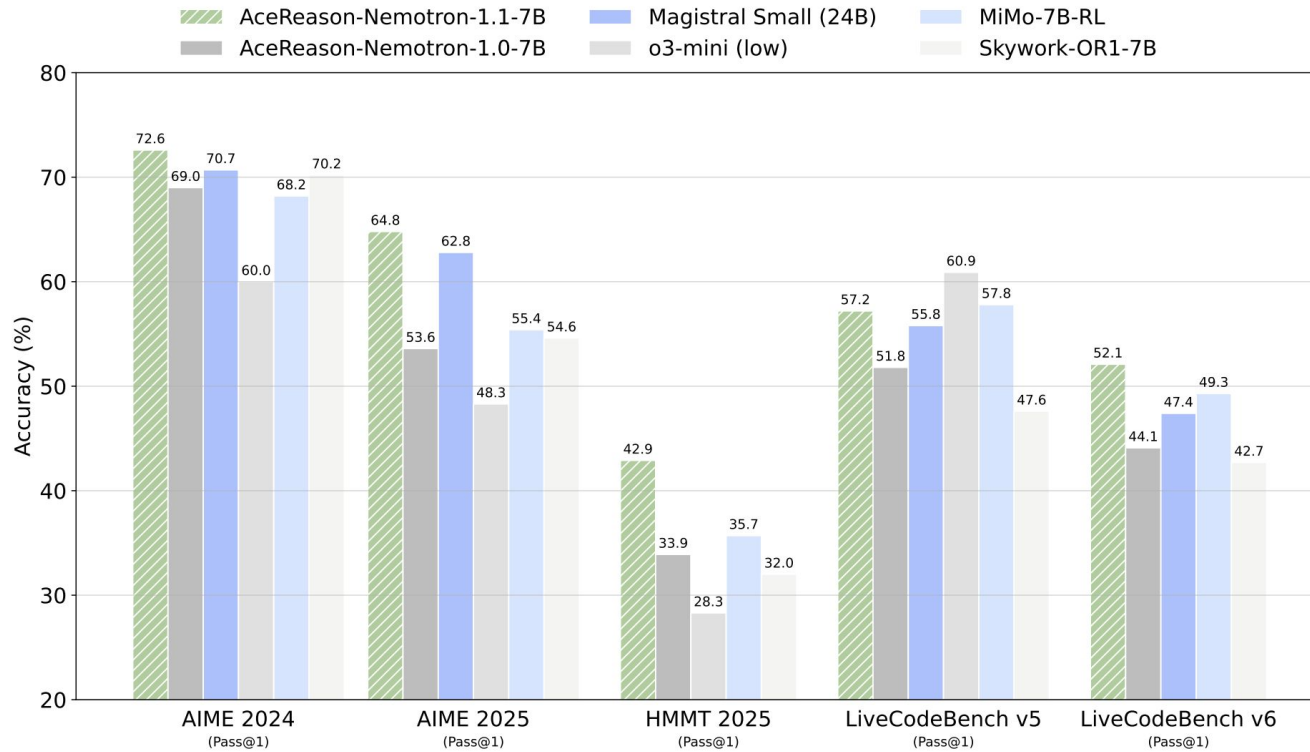


Figure 1: Benchmark accuracy of AceReason-Nemotron-1.1-7B on AIME 2024/2025 (avg@64), HMMT 2025 (avg@64), LiveCodeBench v5 (2024/08/01-2025/02/01, avg@8), and v6 (2025/02/01-2025/05/01, avg@8) using 32,768 output length.

- Different from AceReason-Nemotron-1.0, AceReason-Nemotron-1.1 starts from a stronger SFT model
- As of June 16, 2025, AceReason-Nemotron-1.1 is the leading open-source 7B model in its class.

# Feel free to drop by our poster session!

- Poster ID: **10010287** (Fri, Apr 24, 2026. 10.30 AM to 1 PM local time)
- Paper: <https://arxiv.org/pdf/2506.13284>
- We also open source model weights and data
  - Model weights: <https://huggingface.co/nvidia/AceReason-Nemotron-1.1-7B>
  - Data
    - <https://huggingface.co/datasets/nvidia/AceReason-1.1-SFT>
    - <https://huggingface.co/datasets/nvidia/AceReason-Math>