A. MODEL SUMMARY

A1. Background on you/your team

- Competition Name: AI Mathematical Olympiad - Progress Prize 2
- Team Name: imagination-research
- Private Leaderboard Score: 31/50
- Private Leaderboard Place: 2/2212
- Name: Yichen You
- Location: Beijing, China
- Email: yiyouyc22yyc@gmai.com
- Name: Xuefei Ning
- Location: Beijing, China
- Email: foxdoraame@gmail.com
- Name: Zinan Lin
- Location: Redmond, WA, USA
- Email: linzinan1995@gmail.com

A2. Background on you/your team

- Name: Yichen You

  - **"What is your academic/professional background?"** : I'm an undergraduate student at Tsinghua University.

- **"Did you have any prior experience that helped you succeed in this competition?"** : In my first two years of undergraduate life, I attended some competitions held by my school like software design competition、 hardware design competition、 electronic design Competition etc. Before this competition, I did some general research on MoE models and reasoning models.
- **"What made you decide to enter this competition?"**: I was just interested in everything in llm field including reasoning. We happened to see the Aimo competition being held, and we all thought it was a good opportunity to explore reasoning models.
- **"How much time did you spend on the competition?"**:In the first two weeks, I did some gereral research on reasoning (including learning from the winners' solutions in aimo1) under the guidance of my advisor Xuefei Ning and her friend Zinan Lin. We mainly started from around 1.20, I spent almost all the time on this competition since then. After the competition, I spent several days cleaning our codes and writing the write-up with my advisor Xuefei.
- **"If part of a team, how did you decide to team up?"**: By the end of 2024, Xuefei and Zinan invited me to explore llm's reasoning abality，I was interested in everything in llm field at that time so I agreed immediately.
- **"If you competed as part of a team, who did what?"**: Before we started, I summarized the solutions of the winners in aimo1. In the first 4 weeks, I tried to improve the reasoning efficiency, including the use of lmdeploy and early-stop strategy. Then I spent 4 weeks training the 14b model, including GRPO、 SFT and DPO. In the last 10 days, I added some more detailed control strategies in our solution and helped re-writing the code, organizing the repository under the guidance of Xuefei. After the competition, I helped writing the write-up and slides.

- Name: Xuefei Ning

  - **"What is your academic/professional background?"**: I am now a research-track assistant professor at Tsinghua University.
  - **"Did you have any prior experience that helped you succeed in this competition?"**: I have some experiences in attending and winning some competitions, e.g., serving as the main code&exp contributor and project leader, I got 2nd / 339 teams in NeurIPS'18's adversarial robustness

competition with one collaborator. Besides, I have published papers and independently instructed students to publish papers where idea proposal, problem analysis, solution finding, controlled experimentation, and solid take-away summary are necessary.

- **"What made you decide to enter this competition?"**: Although my "publishing records" are mainly on efficient deep learning techniques, in August 2023, I start pondering around the "reasoning" ability of AI because I am interested in understanding and pushing forwards strong reasoning & continual learning of AI. And my mind experiment somehow "tells" me a strong reasoning ability is needed or at least very useful for data-efficient continual learning. Therefore, the first step is to understand and improve the reasoning ability. In NeurIPS 2024, I have already led and published a NeurIPS paper on reasoning as the first step, in which I get some experience and sense on how the contemporary models behave in math & coding tasks. I want to move faster, look deeper into reasoning and learning, so I think I need to advise a student to do this together with me, as I myself have a lot of other mentoring, teaching, and management responsibilities, and thus cannot continually focus on one project. As such, upon strategical consideration, I think attending a competition will serve as the next good sandbox because it has a clear goal and constraints. By making winning a competition a surrogate goal, a new student without research-oriented judgement & taste & experience can be clearly driven, and we together can get solid experiences on the reasoning topic, touching important training and inference techniques.
- **"How much time did you spend on the competition?"**: After two to three meetings in decision making and rule-checking trials, we mainly start from around 25/01/20. From 01/20, I spent the first 8 weeks mainly advising Yichen on problem analysis (one meeting per week), and the final 10 days rewriting the code, writing analysis scripts and doing ablation studies, and trying out new techniques. And after the announcement, I spent 2~3 days writing the write-up, making the slide, cleaning and open-source the final code, and answering questions on the Kaggle Forum.
- **"If part of a team, how did you decide to team up?"**: Zinan and I are long-time friends and collaborators, I always discuss ideas & thoughts with him, and often invite him to co-advise on lots of projects. Yichen is my advised student, who newly joined my team in late 2024. I find him highly motivated and more importantly, he has a good ability to

understand concepts and logic, has awareness and some ability to do problem analysis, as well as the ability to develop an implementation on his own. Therefore, according to the judgement, I think he is more than suitable for attending a competition as the main code and experiment contributor in our team. BTW, during the competition, I also find Yichen has the habit of proactively engaging with other competitors or open-source project leaders to push forward our project, which is even a characteristic I don't have that much, and has benefited our team a lot.

- **"If you competed as part of a team, who did what?"**: In the first 8 weeks: Problem analysis, proposing conjecture & solutions, deciding the priority together with the team. The whole last 10 days (10~16 hours per day): Re-writing the code, writing analysis and plotting scripts, doing controllable and multi-seed ablation study of multiple techniques (prompt list, quantization, earlystop, etc.), working on improving quantization (RepKV+KV4, QAT)/LLM aggregation/trying to auto-compress the rationale using a non-reasoning model. 2~3 days after announcement, writing the write-up and slide.

- Name: Zinan Lin

  - **"What is your academic/professional background?"** I am a senior researcher at Microsoft Research, working on the fundamentals and applications of generative models (e.g., diffusion models, LLMs).

  - **"Did you have any prior experience that helped you succeed in this competition?"** (1) Research Topics: I co-led, alongside Xuefei, a reasoning-focused paper accepted at NeurIPS 2024 titled "Can LLMs Learn by Teaching for Better Reasoning? A Preliminary Study", which introduces a novel "learning by teaching" paradigm to improve the reasoning capabilities of LLMs in math and coding. In addition, I have led and contributed to several past and ongoing research projects on generating synthetic text and code data using LLMs--areas that are relevant to this competition. (2) Project Execution: Beyond serving as the main or sole contributor on multiple research projects and directions, I also independently advise or co-advise students across various research efforts, guiding them through the entire research lifecycle—from idea and problem formulation to execution, completion, and productization.

  - **"What made you decide to enter this competition?"** One end goal of my research venture is to design AI principles that enable models to learn

and act more like humans. The main long-term direction I am pursuing is a "bottom-up approach", where I rethink and redesign the fundamental formulation of AI models. In parallel, a "top-down approach", where I work with existing models to identify and address their weaknesses, is also important, as it helps me better understand current bottlenecks and guides future model design. Towards this goal, as Xuefei noted in her response (which I fully agree with and support), reasoning remains one of the most significant challenges for LLMs, and participating in a competition is an effective strategy for advancing our understanding in this area. I saw the AIMO competition as an excellent opportunity, given its well-structured organization, challenging problem sets, rigorous evaluation process, and ideal timing (our initial discussion took place in late October 2024, right when the competition was starting). As a result, we decided to form a team and take part in the competition.

- **"How much time did you spend on the competition?"** Our initial discussion took place in October 2024. About a month later, we began holding regular weekly meetings, during which I co-advised Yichen on analyzing results and proposing solutions. Much credit goes to my outstanding teammates, who took on the most time-consuming aspects of the project, particularly on the implementation side.

- **"If part of a team, how did you decide to team up?"** Xuefei and I are long-time friends and collaborators, and we share many common research interests and visions. She invited me to co-advise her student, Yichen. I've been very impressed by Yichen's research capabilities and dedication, and I'm proud of what he has achieved. None of this would have been possible without the efforts of both Yichen and Xuefei.

- **"If you competed as part of a team, who did what?"** I work closely with the team to analyze the math questions, examine failure cases, propose solutions, and plan the overall approach and strategy. The actual coding and implementation are carried out by my amazing teammates.

A3. Summary

- **"The training method(s) you used"**: We use deepseek-r1-distill-qwen-14b as the base model for two-stage training of SFT and DPO
- **"The tool(s) you used"**: We used 360-llama-factory for training

- **"How long it takes to train your model"**: It took 408 A800 hours to train the model

A4. Features Selection / Engineering

- **"External data"**:
  - Evaluation
    - AIME 2025 test set、reference set、HMMT 2025 test set
  - Training
    - light-r1 sft dataset、Limo dataset、open-r1-math-94k

A5. Training Method(s)

- **"What training methods did you use?"**: We used SFT (Supervised Fine-Tuning) and DPO (Direct Preference Optimization) to train our models
- **"Did you ensemble the models?"**: We tried to merge the models obtained from different stages/different datasets, and we did get some models that seemed to have slightly greater performance on the evaluation dataset (AIME), but we did not have enough submission counts to try them.
- **"If you did ensemble, how did you weight the different models?"**: We used the linear method in mergekit to merge the models and tried various merging ratios to get a best one.

A6. Interesting findings

- **"What was the most important trick you used?"**: I think the most important trick is the early-stop strategy in question-level and sample-level, which saves much time for reasoning.
- **"What do you think set you apart from others in the competition?"**: We designed an efficient and complete reasoning framework that can fully utilize the reasoning and coding capabilities of the model; We put much effort on the optimization of reasoning efficiency including weight quantization、and KV cache quantization

A7. Simple Features and Methods

A8. Model Execution Time

- **"How long does it take to train your model?"**: It takes 51 hours to train the final model (on one 8×A800 machine)
- **"How long does it take to generate predictions using your model?"**: It takes 5 hours to generate predictions

A9. References

Wen, L., Xiao, F., He, X., Cai, Y., An, Q., Duan, Z., Du, Y., Liu, J., Tang, L., Lv, X., Zou, H., Deng, Y., Jia, S., & Zhang, X. (2025). Light-R1: Curriculum SFT, DPO and RL for Long COT from Scratch and Beyond. arXiv:2503.10460. https://github.com/Qihoo360/Light-R1

Ye, Y., Huang, Z., Xiao, Y., Chern, E., Xia, S., & Liu, P. (2025). LIMO: Less is More for Reasoning. arXiv:2502.03387 [cs.CL]. https://arxiv.org/abs/2502.03387

open-r1-math-220k. https://huggingface.co/datasets/open-r1/OpenR1-Math-220k

Zou, H., Lv, X., Jia, S., & Zhang, X. (2024). 360-LLaMA-Factory. GitHub. https://github.com/Qihoo360/360-LLaMA-Factory

LMDeploy Contributors. (2023). LMDeploy: A Toolkit for Compressing, Deploying, and Serving LLM. GitHub. https://github.com/InternLM/lmdeploy