

Guiding Through Complexity: What Makes Good Supervision for Hard Math Reasoning Tasks?

Xuan He^{*1}, Da Yin^{*2}, Nanyun (Violet) Peng²¹Tsinghua University, ²University of California, Los Angeles^{*}Equal Contribution

UCLA

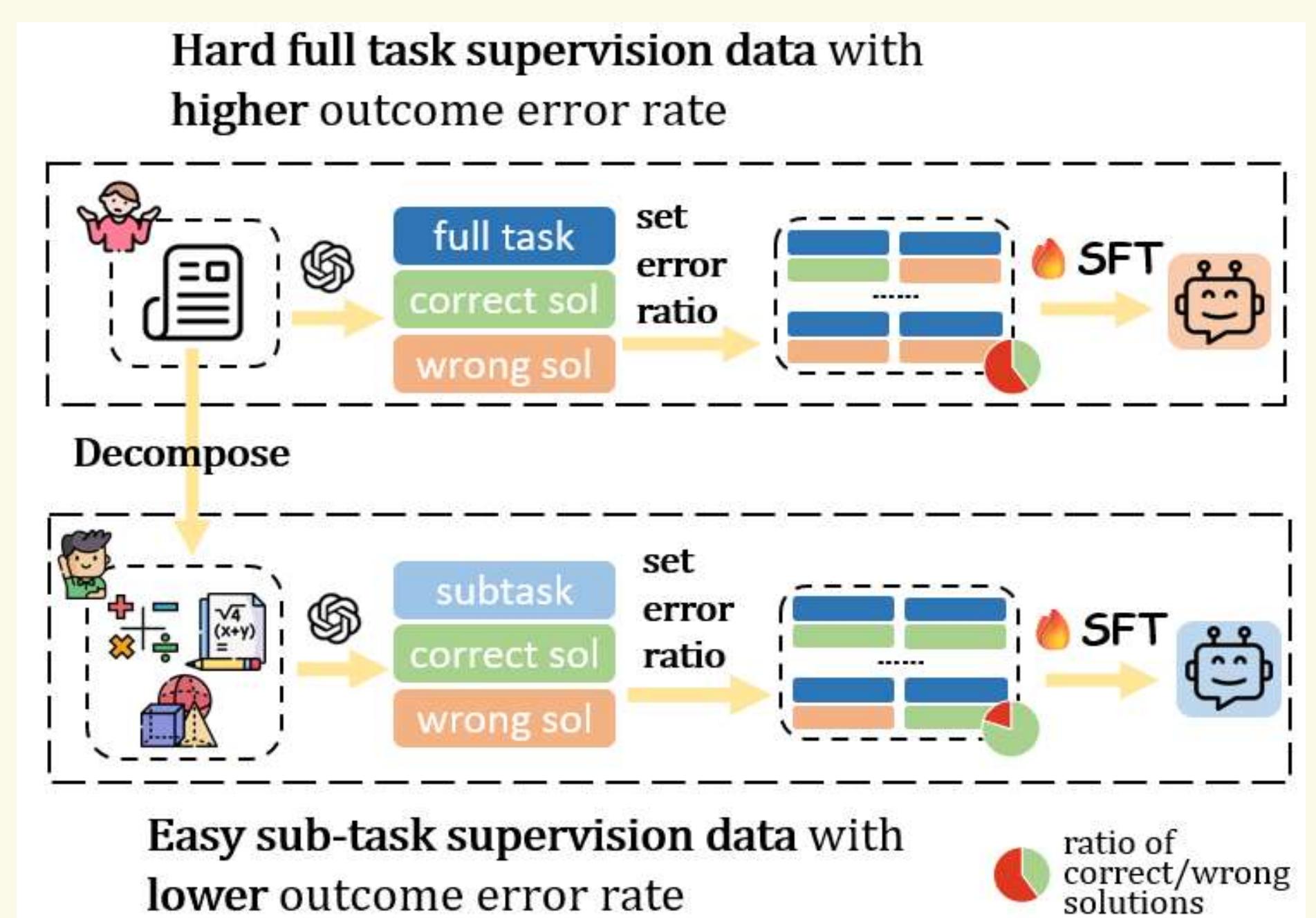
Plus lab



Introduction

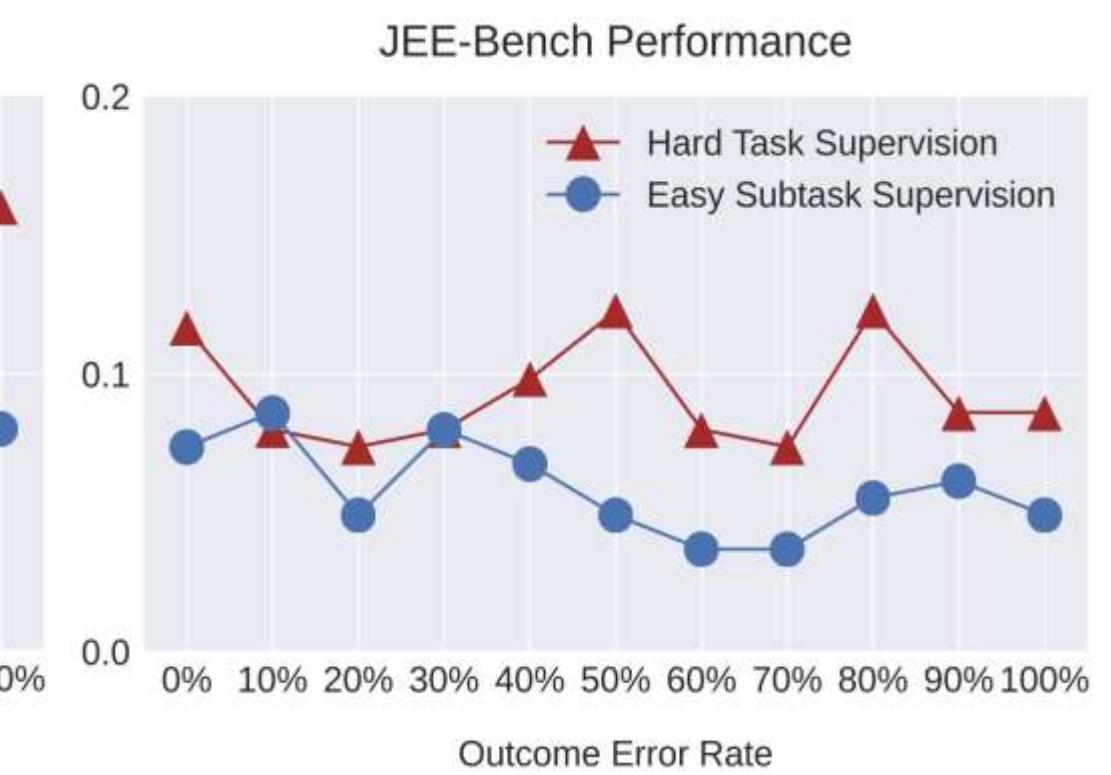
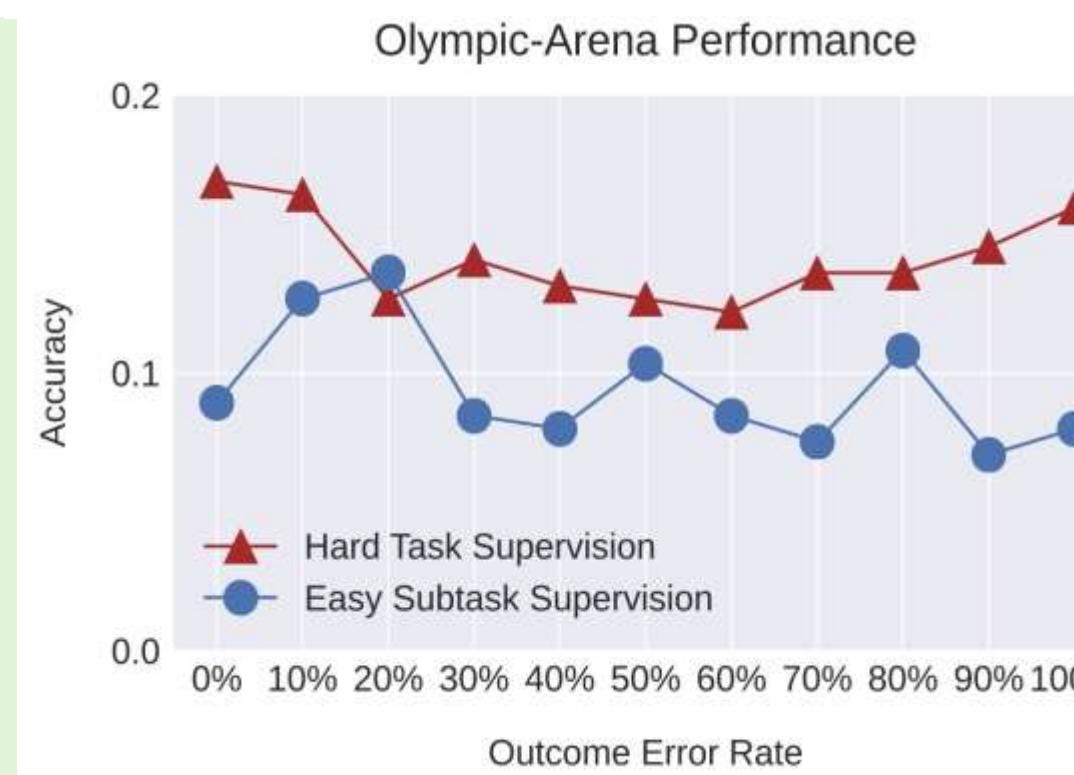
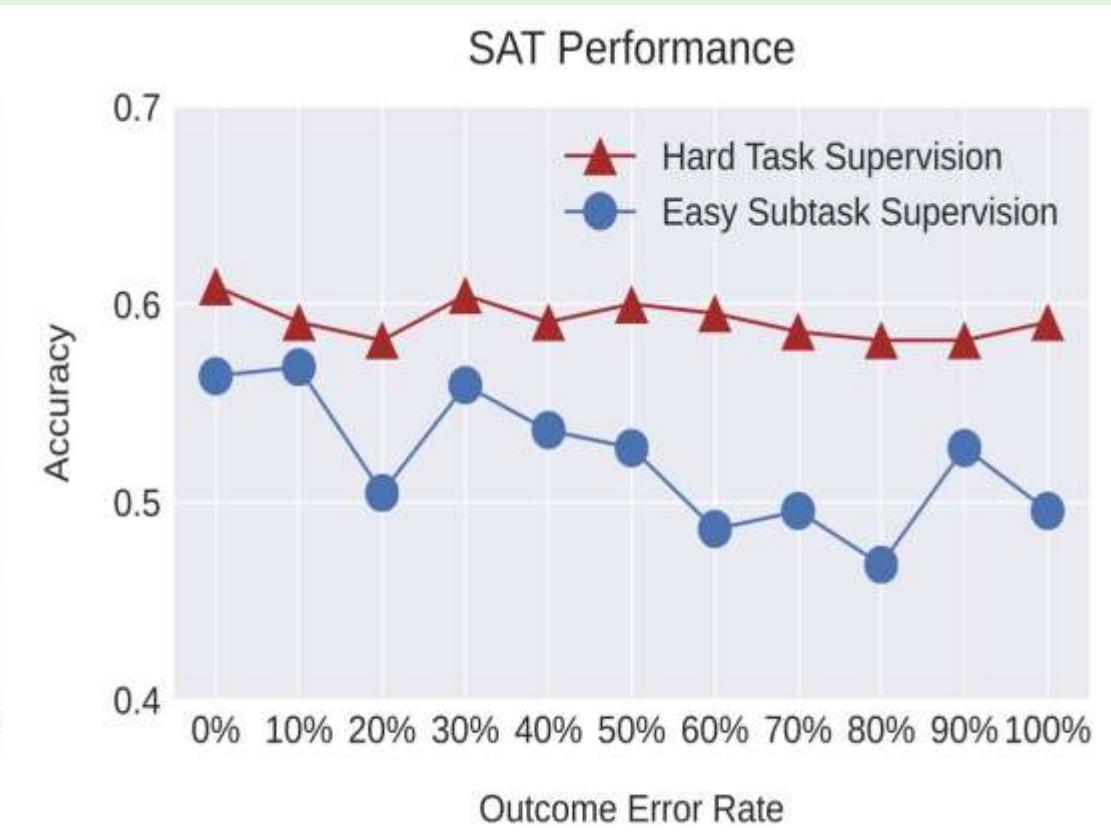
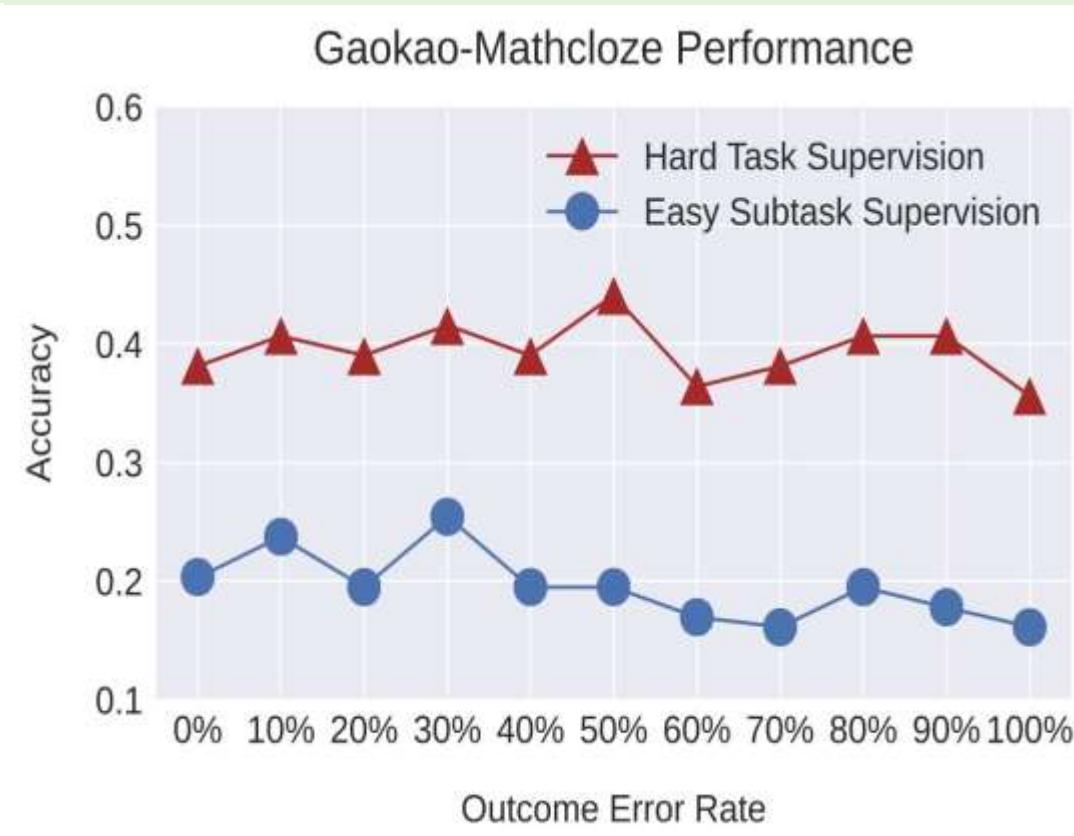
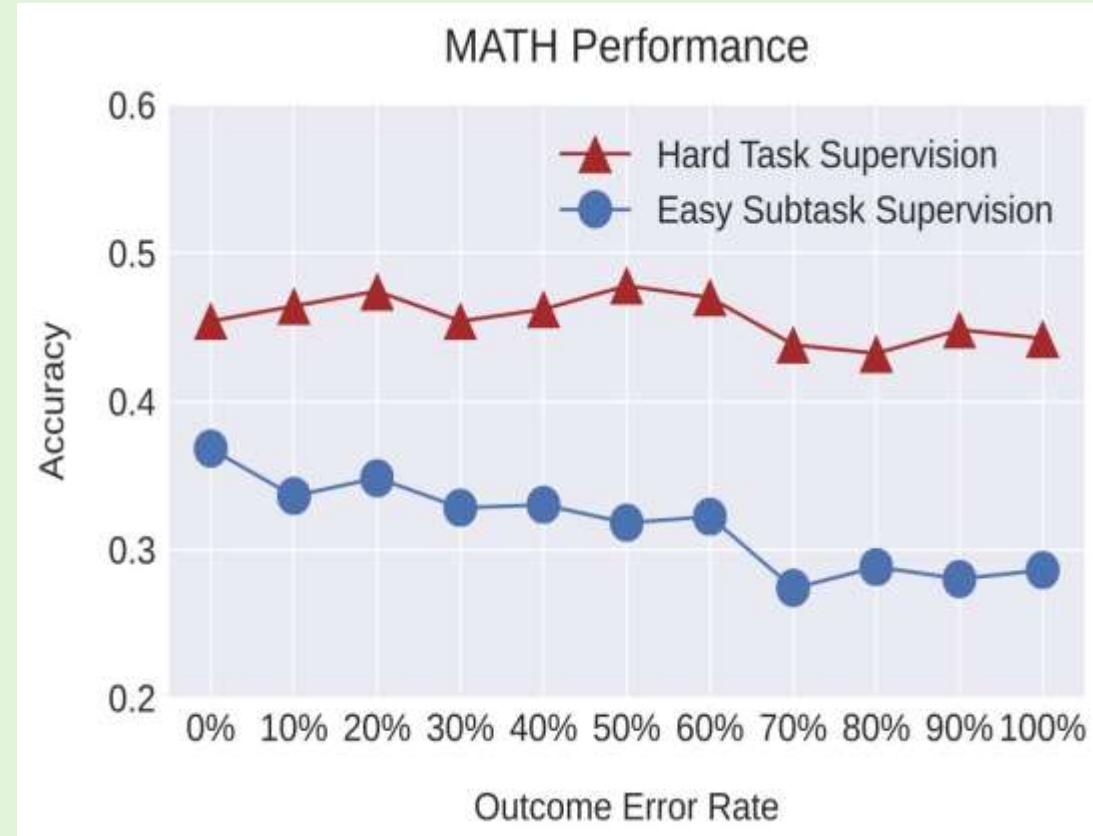
- Exploring data-driven supervision strategies
- “Hard task supervision with low accuracy” outperforms “Easy task supervision with high accuracy”;
- Step-wise error rate matters more than outcome (final result) error rate;
- Combination of hard tasks with decomposed sub-tasks further enhances performance.

New Setting and Task Supervision Synthesis



Which Strategy is Better

Five math benchmarks, from high school level to college competition



Severity of Error/Step-wise ER Matters

Similar outcome error rates with different teacher models.



Stronger teacher models have lower step-wise ER, leading to better performances.

outcome ER 81.8% vs 80.0%

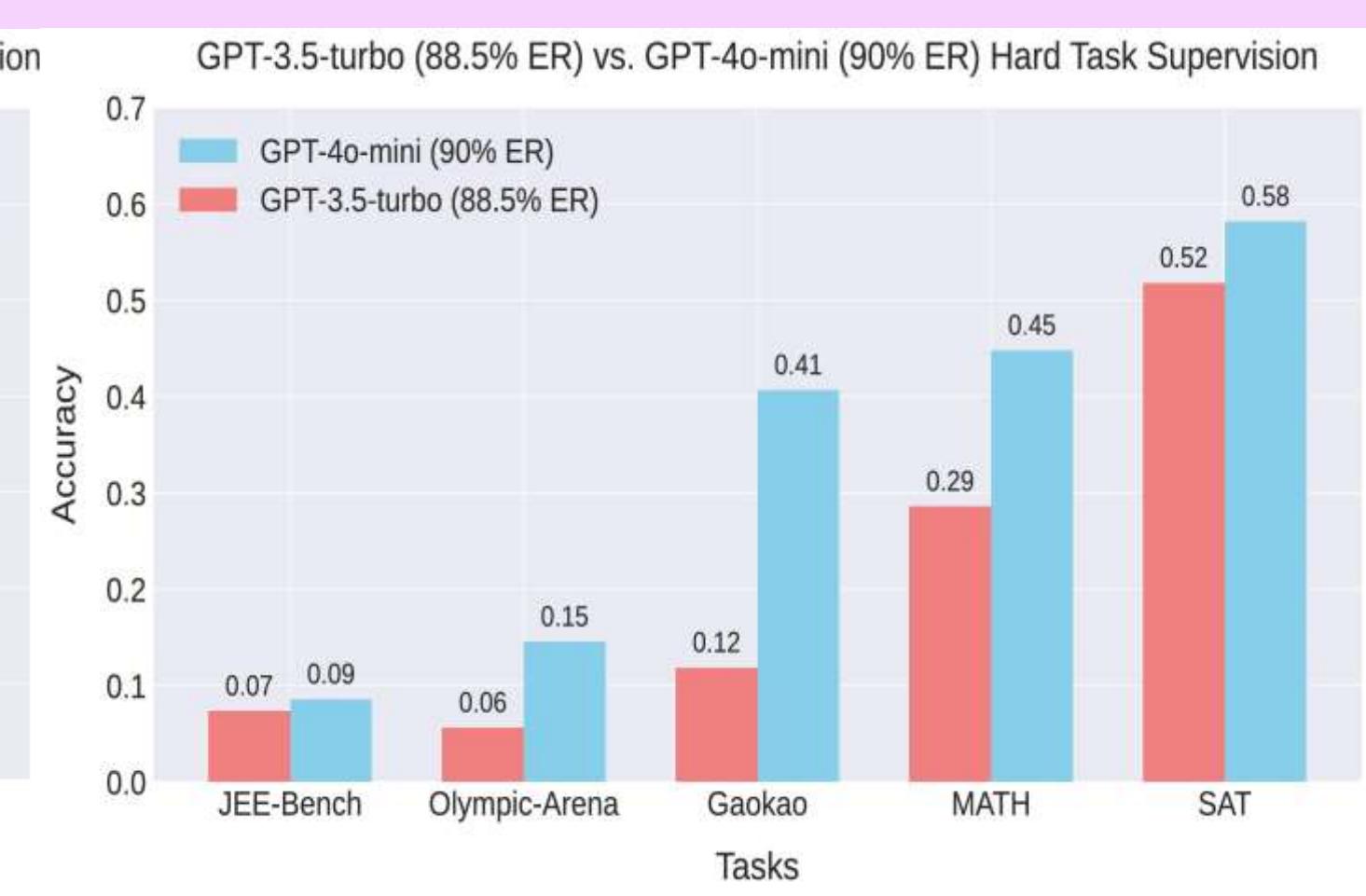
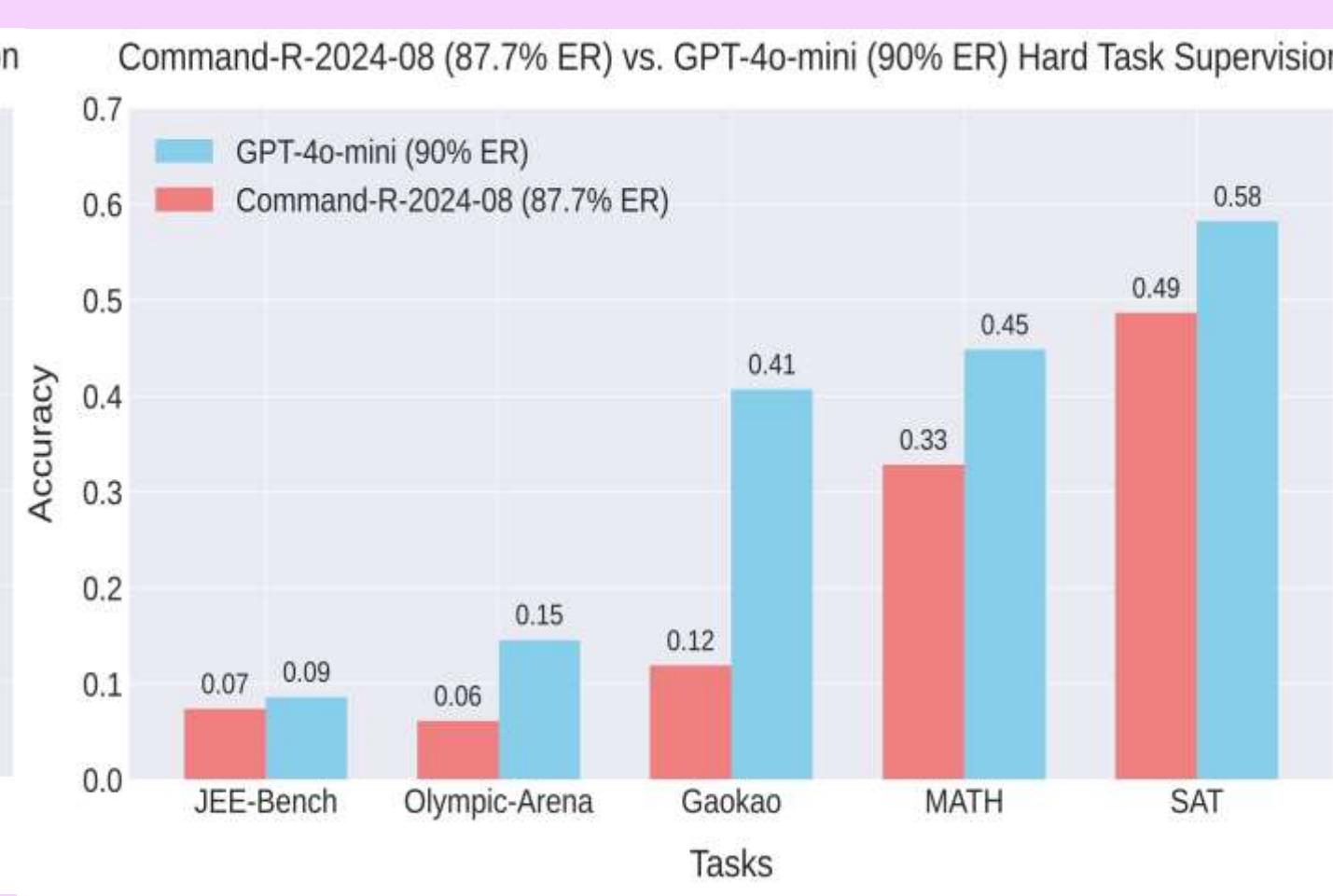
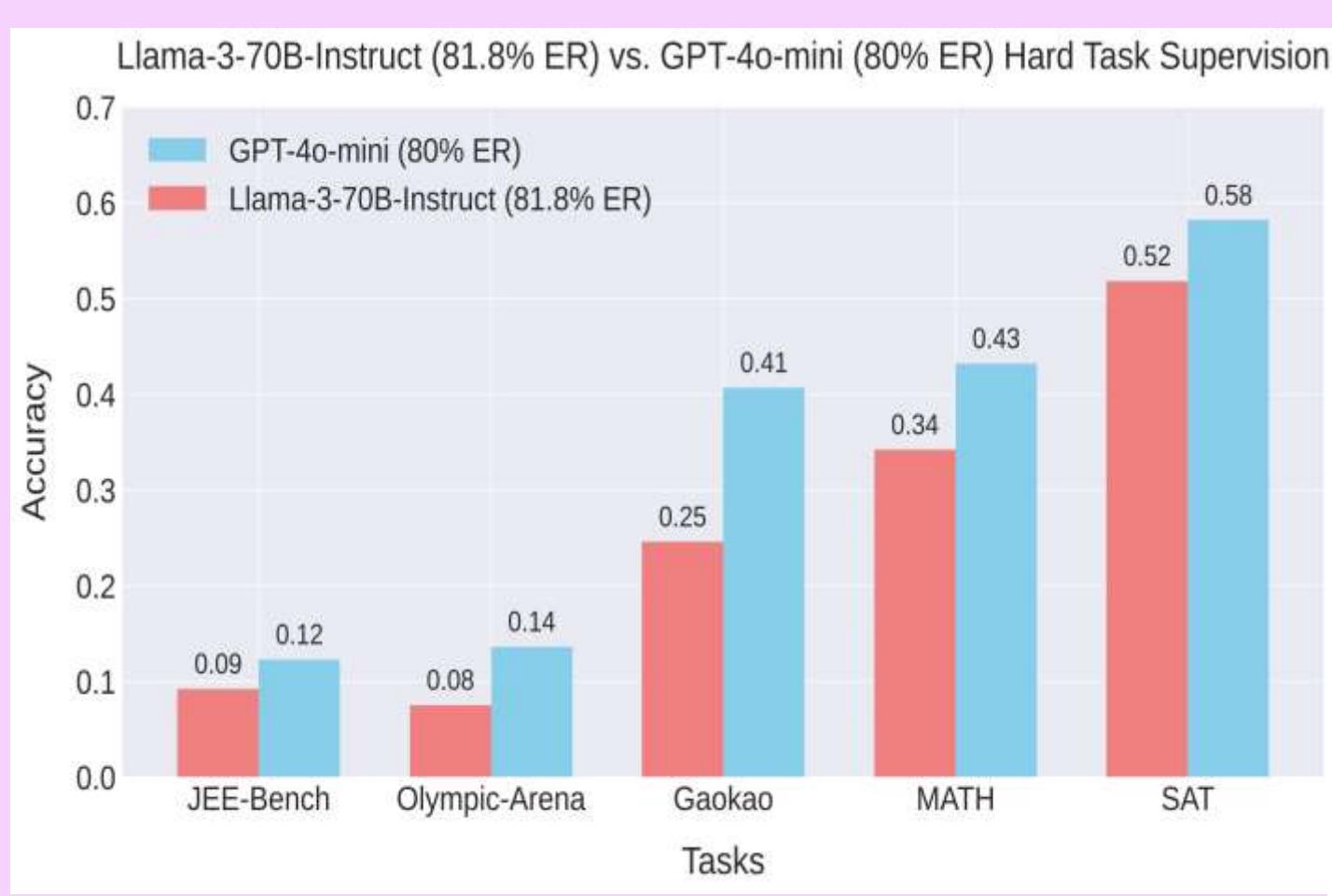
87.7% vs 90.0%

88.5% vs 90.0%

step-wise ER 30.2% vs 11.7%

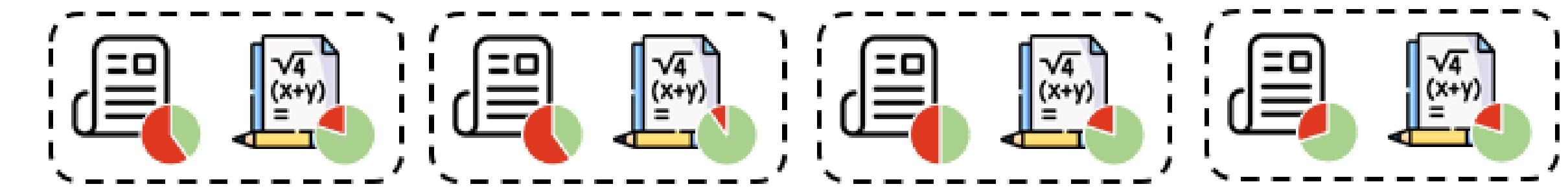
40.0% vs 11.7%

34.7% vs 11.7%



Improvement with only hard supervision: Combination of supervision data

1. Combine with easy subtask supervision



2. Combine with rephrased hard tasks



3. Double training epochs

