

This is a copy of a conversation between ChatGPT & Anonymous.

### Report conversation

You are given W&B logs from several reinforcement learning runs of LLMs at <https://wandb.ai/mbzuai-llm/Guru?nw=nwuserz6cheng>. Identify the number of validation benchmarks in "val" section, where clip-high=0.28, a method adopted in the RL for reasoning paper DAPO, can provide clear performance gain over baseline (w/o clip-high=0.28) on 32B model.

**\*\*Clear Gain Definition:\*\***

We assume "clear gain" on one benchmark if given run A and B:

$\text{max\_validation\_score}(\text{run\_A}) - \text{max\_validation\_score}(\text{run\_B})$

Where run A uses clip-high=0.28 and run B is the baseline without clip-high.

**\*\*Output Format:\*\***

<answer>5</answer>

Answer with the number of validation benchmarks satisfying the requirement (single integer).

To accurately evaluate the number of benchmarks where clip-high=0.28 yields a clear performance gain, could you please confirm the following:

1. Are there specific run names or tags in the W&B logs that identify which runs use clip-high=0.28 and which are baselines?
2. Should we only consider runs on the 32B model size (excluding all others like 7B or 13B)?

3. Is there a fixed list of validation benchmarks you want me to evaluate, or should I consider all available benchmarks in the 'val' section?

Once I have this, I'll start the analysis.

Please refer to the context

Got it. I'll review the W&B logs for the 32B model to identify how many validation benchmarks show a clear performance gain (greater than 0.02) when using clip-high=0.28 compared to the baseline without it.

I'll let you know once I have the results. Feel free to ask anything else in the meantime.



<answer>5</answer>

Sources

