

Instruction on Project Code

We provide the bash script of running ED-iDPO for GSM8K, Math, and s1K tasks under the `scripts/` directory.

- For **GSM8K** and **Math**, we use LLaMA3-8B-Instruct and Qwen2.5-7B-Instruct as the base models, respectively.
- For **s1K**, users should first fine-tune Qwen2.5-7B-Instruct on the s1K training set by following the instructions in [this script](#). The resulting checkpoint should then be used as the base model for ED-iDPO.

The training scripts are located at `scripts/run_[gsm8k|math|s1k].sh`, and the evaluation scripts can be found at `scripts/eval_[gsm8k|math|s1k].sh`, respectively.

To evaluate on the **s1K** out-of-distribution datasets (AIME24, AIME25, MATH500, Minerva Math, Olympiad Bench), users must provide an `openai_key`, as the evaluation relies on an LLM-as-a-judge to assess the correctness of the final answers.

The evaluation script requires the following input arguments:

- `--model`: Path to the model checkpoint
- `--gpu_id`: GPU ID to use for computation
- `--local_port`: Port number to allocate for the local server
- `--accelerate_main_port`: Port number used by the accelerate launcher
- `--output_suffix`: Alias for the current process (useful when running multiple processes in parallel)
- `--current_path`: Root path from which the current process is launched