



Baseline



SFT



GRPO



Step-by-step