



—×— Baseline

—●— SFT

—■— GRPO

—▲— GRPO-8E

—+— GRPO-L