

Qwen3-1.8B: Post-Training Progression (Base → SFT → DPO)

