

1e-4

Qwen3 4B: KL loss during training

