

Pythia 70M: 264 docs used per checkpoint

Δ NLL relative to value at context position 500

