

original_paper:overtrain

~~2.3~~

overtrain rpj-open_lm_1b-32.0

Loss

2.2

10^{11}

10^{12}

Tokens seen