

Perplexity

30
25
20
15

10^8

$10^{8.5}$

10^9

$10^{9.5}$

Number of Model Parameters

Model



Pythia



Mamba



RWKV

