

Cerebras (Param. and Data Scaling)
Pythia 12B Param. (Data Scaling)
Pythia 300B Tokens (Param. Scaling)

Model Family