

Code Pretraining

4K window
1.8T tokens



Long-context Pretraining

16K window
200B tokens



Instruction Fine-tuning

16K window
2B tokens