

Pythia 2.8B: 10,235,000 tokens randomly sampled per token position bin

