1.7B model scale, 300B tokens. Performance across evaluation tasks.

Legend: Nemotron-cc-hq, DCLM, HPLT-2.0, FineWeb-Edu-1.4T, Pile, SlimPajama, CommonCorpus, C4