

Pretraining Composer	inproject				infile			
	FL-4K	PD-4K	PD-16K	Or-16K	FL-4K	PD-4K	PD-16K	Or-16K
Base model (no training)								
$\theta = 10,000$	26.4	36.6	0.0	—	32.6	38.2	0.0	—
$\theta = 500,000$	13.5	16.6	9.8	—	15.5	12.9	4.5	—
File-level 4K								
$\theta = 10,000$	26.2	36.4	0.0	26.2	32.7	38.1	0.0	32.7
$\theta = 500,000$	25.9	36.1	45.2	25.9	33.0	38.1	44.6	33.0
Path Distance .py								
<i>reversed</i>	26.2	37.0	48.8	48.8	33.1	38.7	47.6	48.8
<i>irrelevant</i>	26.1	36.9	48.3	43.2	32.9	38.8	47.5	44.0
	25.8	36.5	47.9	26.7	32.5	38.1	46.7	33.4
Lines IoU .py								
<i>reversed</i>	25.7	36.3	48.7	51.8	33.2	38.4	47.7	50.1
<i>irrelevant</i>	26.1	36.8	48.4	43.5	33.2	38.9	47.4	44.6
	25.8	36.4	47.5	26.7	32.7	38.4	46.6	33.4
Code Chunks .py								
<i>reversed</i>	25.9	36.5	47.9	47.8	32.8	38.2	47.5	47.9
<i>irrelevant</i>	26.1	36.5	47.8	41.3	32.8	38.3	47.4	43.0
	25.8	36.5	47.7	26.9	32.3	37.9	46.3	33.2
Half-memory .py								
<i>reversed</i>	25.7	36.0	47.4	38.6	32.9	38.4	46.6	38.7
<i>irrelevant</i>	25.7	36.2	47.3	35.0	32.9	38.2	46.5	36.9
	25.8	36.0	47.0	27.5	32.4	37.7	46.5	33.0
Declarations .py								
<i>reversed</i>	25.9	36.5	46.8	28.2	32.6	38.1	46.1	34.4
<i>irrelevant</i>	25.7	36.5	46.9	28.1	32.7	38.1	45.7	34.2
	26.2	36.3	47.2	28.2	32.4	38.4	45.6	33.9
Text Chunks .py								
<i>reversed</i>	26.1	36.6	47.5	26.9	33.0	38.5	46.9	33.2
<i>irrelevant</i>	26.0	36.1	47.4	26.8	32.9	38.5	46.2	33.5
	25.9	36.4	47.2	26.8	32.7	38.5	46.2	33.8
Text files								
<i>reversed</i>	25.9	36.2	47.1	26.9	33.0	38.6	46.4	33.5
<i>irrelevant</i>	26.0	36.5	46.9	26.7	33.2	38.4	46.2	33.1
	26.0	36.5	47.1	27.0	32.7	38.2	46.3	33.7
Random files								
	26.2	37.0	48.1	29.8	32.8	38.3	47.3	34.2
Random .py								
	25.9	36.8	48.4	31.9	32.8	38.1	47.0	35.3
Mixed context								
	26.2	36.7	48.5	31.0	32.6	38.2	47.5	36.4
Random tokens								
	26.0	36.2	44.5	26.0	32.6	37.9	45.1	33.1
Duplication								
	19.6	28.8	34.7	96.7	24.5	27.0	28.1	95.0
Leak								
<i>reversed</i>	24.9	34.8	46.1	82.9	30.8	35.5	43.6	81.6
<i>irrelevant</i>	24.3	34.7	45.6	83.8	30.8	35.3	42.8	81.0
	24.5	34.6	45.6	82.2	31.3	35.6	43.2	79.7
Masked Leak								
	25.2	35.4	46.4	65.5	31.6	36.9	45.0	63.5