Open LLM Leaderboard v1

	MMLU	ARC	HellaSwag	Winogrande	TruthfulQA	GSM8K	Average
FLOPs (shared intercept)	9.4	6.2	7.1	5.2	6.4	12.4	7.8
FLOPs	6.5	3.0	**3.0**	2.2	9.0	5.7	4.9
Size and Tokens	7.1	3.2	3.8	2.6	4.5	**5.2**	4.4
PCA + FLOPs (d=1)	8.7	5.6	7.5	4.7	4.5	13.1	7.3
PCA + FLOPs (d=2)	8.7	3.1	4.2	2.1	4.3	13.8	6.0
PCA + FLOPs (d=3)	10.0	3.2	4.8	**2.0**	6.0	12.1	6.3
PCA + FLOPs (d=4)	9.7	3.2	4.9	**2.0**	6.2	12.2	6.4
Sloth basic (d=1)	8.3	6.8	9.8	5.3	6.5	13.5	8.4
Sloth basic (d=2)	7.4	4.1	6.8	3.5	4.2	9.1	5.8
Sloth basic (d=3)	5.8	3.9	6.1	3.4	5.4	7.4	5.3
Sloth basic (d=4)	5.7	4.6	6.3	3.2	5.0	5.9	5.1
Sloth (d=1)	8.2	3.1	4.4	2.5	4.4	11.0	5.6
Sloth (d=2)	**4.4**	**2.7**	4.2	2.9	**3.8**	7.7	4.3
Sloth (d=3)	4.8	2.8	4.0	2.8	4.9	6.3	4.2
Sloth (d=4)	5.3	**2.7**	3.6	2.6	4.2	5.9	**4.1**