Open LLM Leaderboard v1/v2

	MMLU	ARC	HellaSwag	Winogrande	TruthfulQA	GSM8K	IFEval	BBH	MATH Lvl 5	GPQA	MUSR	MMLU-PRO	Average
FLOPs (shared intercept)	9.7	5.1	5.4	4.3	6.6	11.4	7.9	5.3	4.8	2.7	2.6	7.5	6.1
FLOPs	5.6	3.9	**2.6**	2.7	7.1	**5.3**	4.2	3.5	2.2	2.2	4.7	5.4	4.1
Size and Tokens	**4.8**	3.7	3.1	2.5	4.8	6.4	4.1	2.2	6.6	2.2	2.6	6.3	4.1
PCA + FLOPs (d=1)	11.5	6.1	7.9	4.8	4.9	14.7	7.2	4.5	3.3	2.1	2.7	5.9	6.3
PCA + FLOPs (d=2)	11.2	5.1	5.2	3.4	5.2	15.3	7.2	4.4	2.9	2.2	2.7	5.9	5.9
PCA + FLOPs (d=3)	11.0	4.5	4.1	2.5	5.3	15.0	5.7	4.4	3.0	2.1	2.8	6.0	5.6
PCA + FLOPs (d=4)	12.0	4.5	4.4	**2.3**	6.3	13.3	**3.9**	5.9	3.1	2.4	2.9	7.5	5.7
Sloth basic (d=1)	11.3	7.0	7.7	5.6	6.6	15.9	8.6	5.0	3.0	2.3	2.4	6.7	6.8
Sloth basic (d=2)	7.9	5.8	6.0	3.9	4.4	8.9	8.4	3.3	2.0	2.2	1.9	4.7	4.9
Sloth basic (d=3)	7.5	5.8	6.1	3.8	4.6	8.6	4.0	3.4	2.3	1.9	2.2	4.3	4.5
Sloth basic (d=4)	6.7	6.0	6.2	3.8	8.4	7.1	**3.9**	4.0	2.3	**1.8**	4.1	4.3	4.9
Sloth (d=1)	8.6	5.1	7.2	5.3	**4.3**	12.3	6.7	2.8	2.9	2.0	1.7	4.4	5.3
Sloth (d=2)	5.8	3.8	5.0	3.3	5.1	7.5	7.2	2.6	2.2	2.2	1.9	3.3	4.2
Sloth (d=3)	5.4	**3.0**	4.4	3.1	5.0	6.4	4.7	2.5	**1.6**	2.2	**1.5**	**2.5**	**3.5**
Sloth (d=4)	5.4	4.0	5.0	2.9	6.5	6.8	4.1	**1.8**	2.8	2.1	3.8	3.8	4.1