Open LLM Leaderboard v1

	MMLU	ARC	HellaSwag	Winogrande	TruthfulQA	GSM8K	Average
FLOPs	9.2	7.9	5.5	5.0	8.8	20.0	9.4
Size and Tokens	7.5	3.1	3.4	**1.9**	5.6	7.9	4.9
PCA + FLOPs (d=1)	10.8	9.0	13.1	7.2	5.5	13.8	9.9
PCA + FLOPs (d=2)	10.8	8.3	12.2	6.0	5.2	12.8	9.2
PCA + FLOPs (d=3)	11.7	8.4	11.6	6.0	5.0	12.2	9.2
PCA + FLOPs (d=4)	12.7	8.3	11.3	5.8	8.1	12.1	9.7
Sloth basic (d=1) (shared intercept)	8.4	6.0	7.2	5.5	6.6	9.7	7.2
Sloth basic (d=2) (shared intercept)	7.6	6.1	6.5	4.7	4.8	8.8	6.4
Sloth basic (d=3) (shared intercept)	7.7	6.2	6.5	4.9	6.4	7.5	6.5
Sloth basic (d=4) (shared intercept)	7.6	6.2	6.6	4.9	6.3	7.3	6.5
Sloth (d=1) (shared intercept)	8.9	5.2	6.9	5.5	7.7	13.1	7.9
Sloth (d=2) (shared intercept)	8.0	5.2	5.9	4.9	6.3	10.9	6.9
Sloth (d=3) (shared intercept)	7.7	5.6	5.8	4.9	5.8	10.8	6.8
Sloth (d=4) (shared intercept)	8.0	5.1	5.7	4.6	5.8	11.5	6.8
Sloth basic (d=1)	11.9	6.4	7.9	4.8	5.3	19.1	9.2
Sloth basic (d=2)	5.3	4.9	6.9	3.3	**4.1**	10.2	5.8
Sloth basic (d=3)	4.9	4.8	6.0	3.1	5.6	8.6	5.5
Sloth basic (d=4)	4.5	5.1	5.9	3.1	5.0	7.4	5.2
Sloth (d=1)	8.5	4.2	5.5	4.0	5.3	14.8	7.0
Sloth (d=2)	**4.0**	**2.8**	3.6	2.2	5.1	**6.9**	**4.1**
Sloth (d=3)	4.7	3.2	3.5	**1.9**	4.3	7.5	4.2
Sloth (d=4)	5.0	3.0	**3.3**	2.0	4.9	8.0	4.3