

#Train models	0.35B (X500)	0.76B (X230)	1.3B (X134)	2.7B (X64)	6.7B (X26)	13.0B (X13)
2	13 [★]	17	2.5 [★]	4.8	6.4	4.7
3		23	10	12	7.7	5.9
4			5.6	10	9.1	6.7
5				5.1	0.54	7.6
6					0.92 [★]	3.1
7						2.7

Largest model parameters (scale up predicted)