Metric	M_1	M_2	Averaged	Gold Model	Aggregate Model	Tuned Aggregate Model
Accuracy	0.51	0.51	0.66	0.92	0.83	0.88
Loss	2.05	1.88	1.20	0.27	0.54	0.39