L2 Distance

0.9
0.8
0.7
0.6

0.54	0.73	0.94	0.91

model

- Hyp. Minds
- No Hyp. Eval.
- ReAct
- Base LLM