Negative Accuracy

	Original	+2	+4	+8	+16
GPT-4 Turbo	100.0	100.0	100.0	99.3	96.5
GPT-3.5 Turbo	86.9	87.6	86.9	86.2	92.4
Llama3 70B	93.6	98.6	97.9	97.9	94.5
Mixtral 8*22B	93.1	96.5	97.9	95.2	91.7
Llama3 8B	86.2	93.1	93.1	95.9	91.0
Mistral 7B	96.5	98.6	97.2	97.2	95.9