

Harmonic Score per Model and Prompt

Model

Zephyr 7B	0.660	0.671	0.537	0.438
LLaMA 3.1 8B	0.498	0.431	0.456	0.613
OpenChat 3.5-0106	0.485	0.507	0.502	0.498
OpenChat 3.5	0.578	0.506	0.462	0.418
InternLM2.5 7B	0.574	0.512	0.373	
OpenChat 3.5 GPTQ	0.570	0.526	0.435	0.402
Nous Hermes Yi 34B	0.445	0.497	0.422	0.386
Qwen2 7B	0.518	0.476	0.341	0.403
Nous Hermes Mistral 7B DPO				0.428
LLaMA 2 13B				0.384
LLaMA 3 8B				0.370
Yi 6B	0.250	0.399	0.386	
Gemma 7B	0.393	0.410	0.007	0.403
LLaMA 2 7B	0.000	0.000	0.000	0.387
GPT-2-Small	0.180	0.000	0.000	
StableBeluga 7B	0.000	0.000	0.004	

prompt1      prompt2      prompt3      prompt4

Prompt

