Claude-Sonnet-A Claude-Sonnet-A Claude-net-3.7	Ratin	g Task -	Spearr	man r	
Claude-Sonnet-A Claude-Sonnet-A Claude-Sonnet-A Claude-Sonnet-A Claude-Sonnet-A Claude-Sonnet-A Claude-Sonnet-A Claude-Sonnet-A Claude-Sonnet-A Claude-Opt-3-7 Claude-Sonnet-A Claude-Opt-3-7 Claude-Sonnet-A Claude-Opt-3-7 Claude-Sonnet-A Claude-Net-3-235B	0.34	0.13	0.21	0.27	- 0.35
clande of 3.	0.32	0.09	0.19	0.27	
Clayon, 35B-	0.33	0.15	0.22	0.30	
claude 23-13-8-	0.32	0.09	0.19	0.27	- 0.30
Claude 3.35B - Claude 3.32B - Claude 3.30B - Claude	0.26	0.12	0.13	0.29	
On Color	0.36	0.24	0.18	0.24	0.25
Claude Se 3.235 Claude Swen3.32B - Qwen3.30B - Qwen3.34B - Qwen3.38B - Qwen3.4B - Qwen3.	0.31	0.11	0.15	0.25	- 0.25
QWED3'1B-	0.33	0.19	0.15	0.18	
QV 2021B -	0.33	0.16	0.05	0.21	- 0.20
ON3-1-6B-	0.04	0.01	0.01	0.02	0.20
ONE.3-0.	0.02	-0.02	-0.01	0.01	
Qwen3-1.6B- Qwen3-0.6B- Qwen3-0.6B-	0.34	0.13	0.21	0.33	- 0.15
(-\\' \	0.33	0.17	0.19	0.28	
GPT-Shano GPT-5-Nini GPT-40-Pro	0.31	0.11	0.12	0.21	
GP, 40-1910 -	0.36	0.12	0.14	0.26	- 0.10
GP1-2:1356-	0.31	0.15	0.17	0.25	
Gemis Bishir	0.30	0.12	0.17	0.27	0.05
cemilific stest.	0.28	0.07	0.08	0.25	- 0.05
wini-	0.27	0.07	0.10	0.27	
ser Grokeeksorie	0.30	0.12	0.14	0.25	- 0.00
Deep Re	0.29	0.09	0.14	0.29	0.00
0.28 0.07 0.08 0.25 -0.05					