Rank Across Categories	All	Real World	Domain Knowledge	Specificity	Technical Accuracy	Problem Solving	Creativity	Complexity	Math
1	1	1	1	1	1	1	1	1	5
2	2	2	2	2	2	2	2	5	1
3	3	3	3	3	3	3	3	3	3
4	4	4	4	5	5	5	4	2	2
5	5	5	5	4	4	7	9	7	7
6	6	6	6	8	6	4	8	8	8
7	7	7	7	6	7	8	6	4	6
8	8	9	9	7	11	9	10	9	9
9	9	8	10	9	8	6	7	12	15
10	10	10	8	13	9	12	11	6	13
11	11	11	11	12	12	11	5	13	16
12	12	12	12	11	14	13	16	11	12
13	13	13	13	10	13	10	13	15	18
14	14	14	14	14	10	16	12	16	4
15	15	16	15	15	15	15	17	14	17
16	16	15	16	16	16	14	14	17	21
17	17	17	17	18	17	17	18	18	14
18	18	18	18	17	18	18	15	23	39
19	19	20	20	20	20	23	20	10	10
20	20	19	19	23	23	21	22	21	11

LLMs (Ranked by All)

1. chatgpt-4o-latest
2. gemini-1.5-pro-exp-0801
3. gpt-4o-2024-05-13
4. gpt-4o-mini-2024-07-18
5. claude-3-5-sonnet-20240620
6. gemini-advanced-0514
7. llama-3.1-405b-instruct
8. gpt-4o-2024-08-06
9. gemini-1.5-pro-api-0514
10. gemini-1.5-pro-api-0409-preview
11. athene-70b-0725
12. mistral-large-2407
13. gpt-4-turbo-2024-04-09
14. llama-3.1-70b-instruct
15. claude-3-opus-20240229
16. gpt-4-1106-preview
17. yi-large-preview
18. gpt-4-0125-preview
19. reka-core-20240722
20. gemini-1.5-flash-api-0514
21. deepseek-v2-api-0628
22. gemma-2-27b-it
23. deepseek-coder-v2-0724
39. deepseek-coder-v2