

| | | Queries and Errors per Model | | |
|-------|--------------------------|------------------------------|---------|-----------------------|
| Model | Claude 3 Haiku | 29730 | 0 | 0 |
| | Claude 3 Sonnet | 27860 | 0 | 8 |
| | Claude 3 Opus | 14640 | 0 | 2 |
| | Code Llama 7B | 35780 | 9 | 1227 |
| | Code Llama 13B | 37170 | 2 | 241 |
| | Code Llama 34B | 40170 | 0 | 6 |
| | Gemini 1.5 Flash | 23980 | 0 | 0 |
| | Gemini 1.5 Pro | 12780 | 0 | 0 |
| | GPT-3.5 Turbo (0613) | 39050 | 0 | 0 |
| | GPT-3.5 Turbo (1106) | 31550 | 0 | 0 |
| | GPT-3.5 Turbo (0125) | 35620 | 0 | 0 |
| | GPT-4 (0314) | 14040 | 0 | 0 |
| | GPT-4 (0613) | 13290 | 0 | 0 |
| | GPT-4 Turbo (1106) | 15165 | 0 | 0 |
| | GPT-4 Turbo (2024-04-09) | 13560 | 0 | 0 |
| | GPT-4o (2024-05-13) | 15225 | 0 | 0 |
| | Llama 2 Chat 7B | 27500 | 0 | 0 |
| | Llama 2 Chat 13B | 29020 | 0 | 0 |
| | Llama 2 Chat 70B | 27230 | 0 | 0 |
| | Llama 3 Instruct 8B | 25570 | 0 | 0 |
| | Llama 3 Instruct 70B | 23730 | 0 | 0 |
| | Mistral 7B | 25390 | 4 | 378 |
| | Mixtral 8x7B | 23650 | 0 | 0 |
| | Phi-2 | 34190 | 383 | 3034 |
| | Yi Chat 34B | 42120 | 1075 | 2446 |
| | | | Queries | Both Responses Errors |