# Accuracy per Domain and Model

| Model | Algebra | Calculus | Combinatorics | Geometry | Mathematical Logic | Probability | Theory of Numbers |
|---|---|---|---|---|---|---|---|
| o3−mini | 89% | 72% | 72% | 75% | 82% | 61% | 89% |
| LLaMA−3.3−70B | 57% | 42% | 30% | 24% | 55% | 36% | 24% |
| Grok−4 | 94% | 71% | 85% | 67% | 81% | 63% | 91% |
| Grok−3 | 74% | 49% | 47% | 34% | 67% | 45% | 55% |
| GPT−4o−mini | 63% | 45% | 32% | 17% | 42% | 37% | 33% |
| GPT−4o | 68% | 51% | 42% | 21% | 54% | 40% | 34% |
| DeepSeek−V3−685B | 87% | 59% | 45% | 37% | 77% | 48% | 79% |
| Claude−Opus−4 | 84% | 59% | 47% | 41% | 75% | 52% | 78% |

Accuracy

100%
75%
50%
25%
0%