

Benchmark: MMLU Elementary Maths  
Performance Metric: Brier Score  
Correlation Metric: Kendall

