

Benchmark: MMLU Virology
Performance Metric: Brier Score
Correlation Metric: Kendall

