

Benchmark: MMLU Human Sexuality
Performance Metric: Brier Score
Correlation Metric: Spearman

