

Benchmark: MMLU Professional Psychology
Performance Metric: Brier Score
Correlation Metric: Spearman

