

Benchmark: MMLUHS Psychology
Performance Metric: Brier Score
Correlation Metric: Spearman

