

Benchmark: MMLUComp Security
Performance Metric: Brier Score
Correlation Metric: Spearman

