

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

