

Benchmark: MMLU HS Euro History
Performance Metric: Brier Score
Correlation Metric: Kendall

