

Benchmark: MMLU International Law
Performance Metric: Brier Score
Correlation Metric: Kendall

