

Benchmark: MMLUProfessional Law
Performance Metric: Brier Score
Correlation Metric: Kendall

