

Benchmark: MMLUMedical Genetics
Performance Metric: Brier Score
Correlation Metric: Kendall

