

Benchmark: MMLU Formal Logic
Performance Metric: Accuracy
Correlation Metric: Kendall

