

Benchmark: MMLU Professional Psychology  
Performance Metric: Accuracy  
Correlation Metric: Kendall

