

Benchmark: MMLU Human Sexuality  
Performance Metric: Accuracy  
Correlation Metric: Spearman

