

Benchmark: MMLU Moral Scenarios
Performance Metric: Accuracy
Correlation Metric: Spearman

