

Benchmark: MMLU World Religions
Performance Metric: Target Greedily Decoded
Correlation Metric: Spearman

