

Benchmark: MMLU Sociology
Performance Metric: Target Greedily Decoded
Correlation Metric: Spearman

