

Benchmark: MMLU Moral Disputes
Performance Metric: Target Greedily Decoded
Correlation Metric: Spearman

