

Benchmark: MMLU Global Facts
Performance Metric: Target Greedily Decoded
Correlation Metric: Kendall

