

Benchmark: MMLU Anatomy
Performance Metric: Target Greedily Decoded
Correlation Metric: Kendall

