

Benchmark: MMLU College Medicine
Performance Metric: Target Greedily Decoded
Correlation Metric: Kendall

