

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

