

Benchmark: MMLU HS Biology
Performance Metric: Target Greedily Decoded
Correlation Metric: Pearson

