

Benchmark: MMLU Business Ethics
Performance Metric: Target Greedily Decoded
Correlation Metric: Spearman

