

Benchmark: MMLU HS Govt & Politics  
Performance Metric: Target Greedily Decoded  
Correlation Metric: Spearman

