

Benchmark: MC-TACO
Performance Metric: Target Greedily Decoded
Correlation Metric: Spearman

