

Benchmark: PIQA
Performance Metric: Target Greedily Decoded
Correlation Metric: Spearman

