

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

