

Benchmark: Social Interaction QA
Performance Metric: Accuracy
Correlation Metric: Spearman

