

Benchmark: OpenBookQA
Performance Metric: Brier Score
Correlation Metric: Pearson

