

Benchmark: ARC-Challenge
Performance Metric: Accuracy
Correlation Metric: Spearman

