

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

