

Benchmark: HellaSwag
Performance Metric: Accuracy (Normalized)
Correlation Metric: Spearman

