

Benchmark: HellaSwag  
Performance Metric: Accuracy  
Correlation Metric: Spearman

