

Benchmark: XWinograd-EN  
Performance Metric: Brier Score  
Correlation Metric: Spearman

