

XWinograd-EN

$p_{\theta}^{\text{Choices}}$ (Correct Choice)

Per-Sample Correlation(Compute, Score) for
Pythia 12B Param. (Data Scaling)

1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00

correlation_metric

- kendall
- pearson
- spearman

Per-Sample Correlation(Compute, Score) for
OLMo 7B Param. (Data Scaling)

-1.00 -0.75 -0.50 -0.25 0.00 0.25 0.50 0.75 1.00

