

HellaSwag

 $\log p_{\theta}^{\text{Vocab}}(\text{Correct Choice})$ Per-Sample Correlation(Compute, Score) for
Pythia 12B Param. (Data Scaling)