

HellaSwag

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

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

