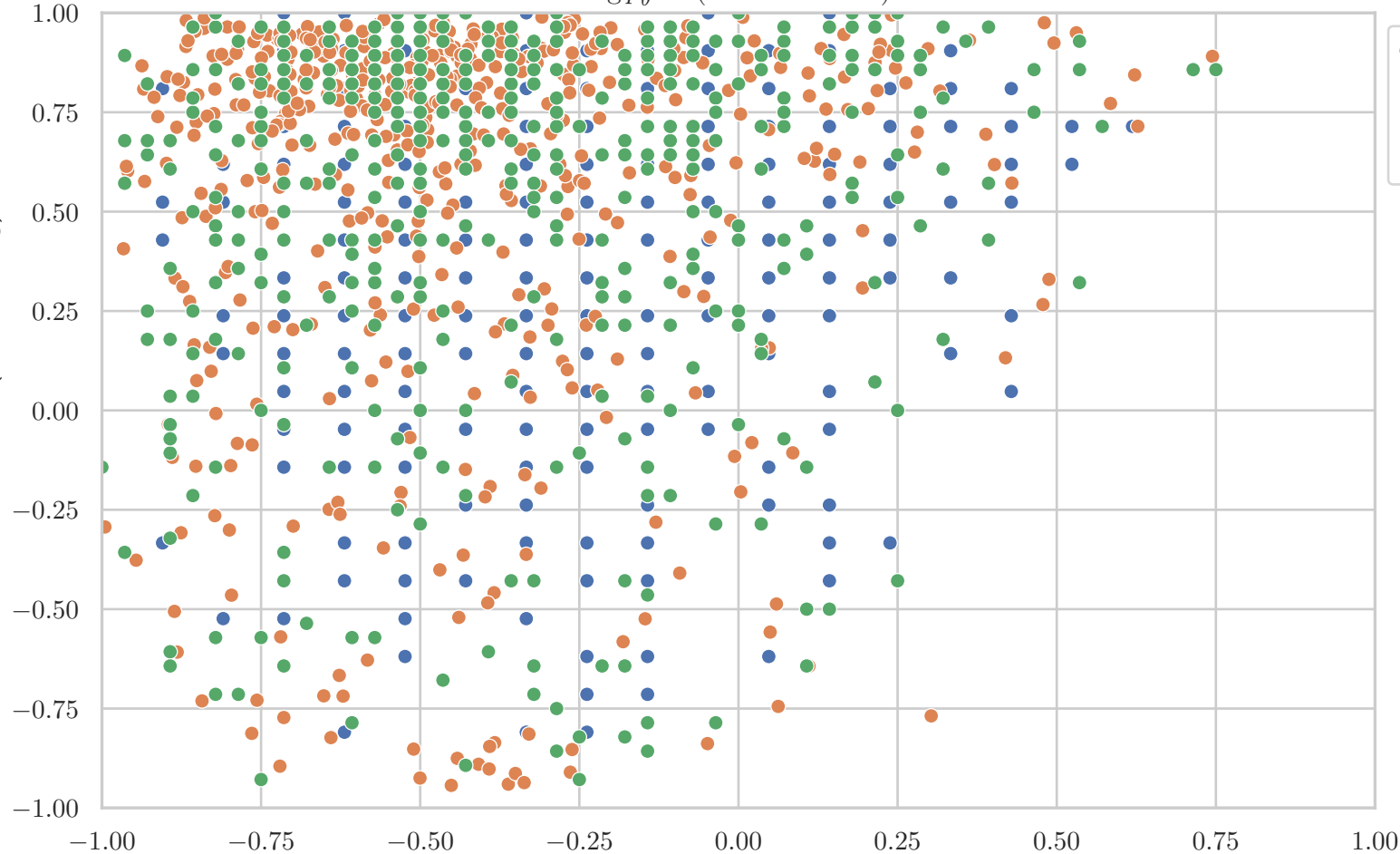


PubMedQA

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

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



correlation_metric

- kendall
- pearson
- spearman

Per-Sample Correlation(Compute, Score) for
Cerebras (Param. and Data Scaling)