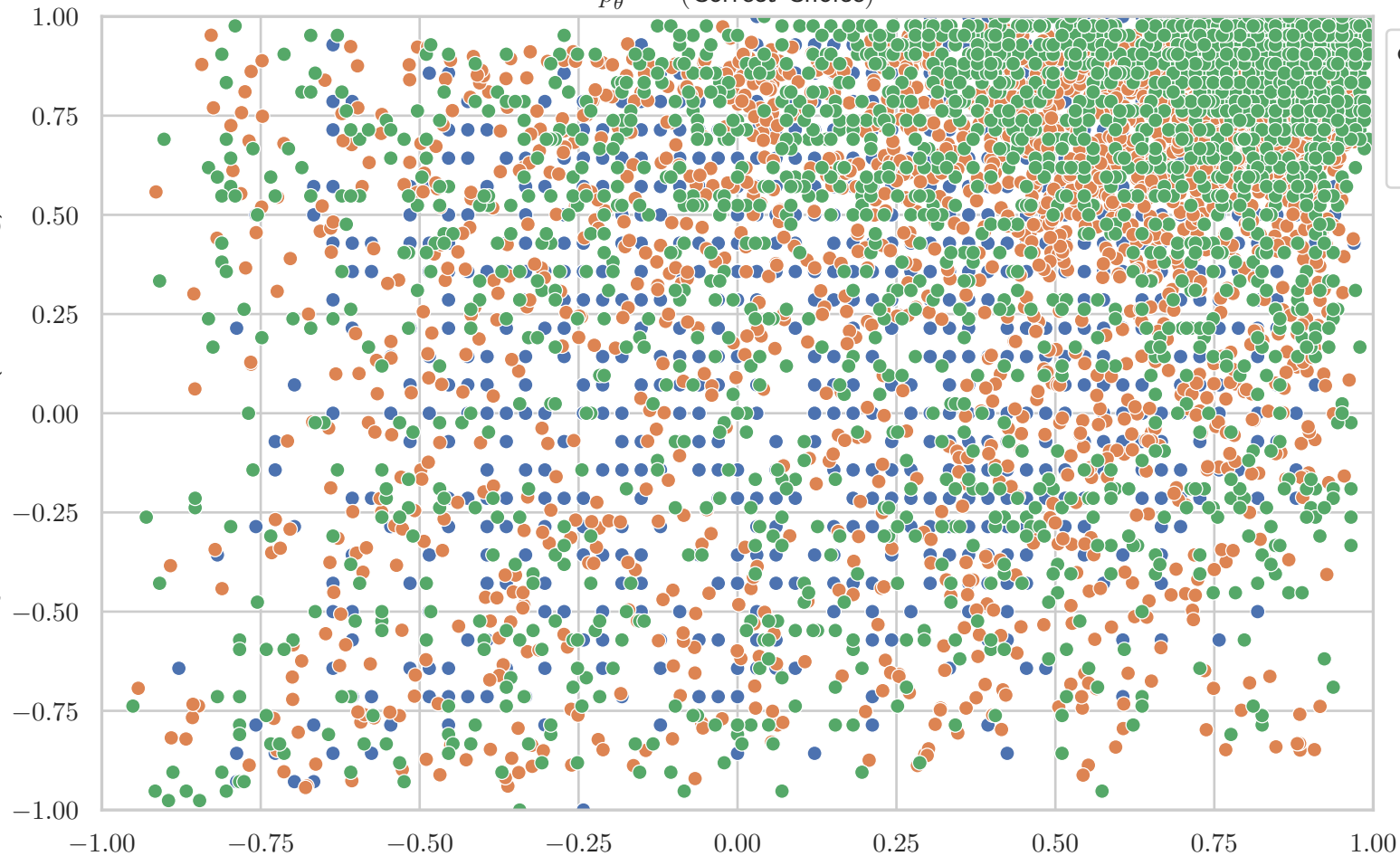


ARC-Easy

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

correlation_metric

kendall

pearson

spearman

Per-Sample Correlation(Compute, Score) for
LLM360 Amber 7B Tokens (Param Scaling)