

MMLU Logical Fallacies

Brier Score

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

