

probes

$$R^2_{34,34} = \overline{\overline{R^2_{34,34}}}$$

$$R^2_{55,28} = \overline{\overline{R^2_{55,28}}}$$

$$R^2_{67,94} = \overline{\overline{R^2_{67,94}}}$$

$$R^2_{74,97} = \overline{\overline{R^2_{74,97}}}$$

Where

is

X

in

the

world?

X

$$R^2$$

square_world_decoders.png Decoder models

square_world_encoders.png Encoder models

$$R^2$$

perf. While we notice that the performance increases consistently for every continent, we do not observe a significant reduction

continent.png Average MSE by continent for different sizes in the Pythia suite.

$(x_i)_{i \in [1, N]}$

$$Gini(x) = \frac{\sum_{i,j \in [1, N]} |x_i - x_j|}{N \cdot \sum_{i=1}^N x_i}$$

combined.png Gini coefficient of MSE on the test set averaged by country or by continent, as model size increases.

10050.png Test log - MSE for Pythia - 1B as plotted on a World map.

1

Fully

fully

size.png Pearson correlation coefficient of various factors with location - wise MSE, for several Pythia model sizes. *: Test

map, we observe that the error on coordinates prediction is negatively correlated with the latitude, i.e. southern locations are less

gender discuss the relevance of scaling model size over larger magnitudes, with regard to environmental and financial cost

17r_n at biblrec - coling2024 - natbib