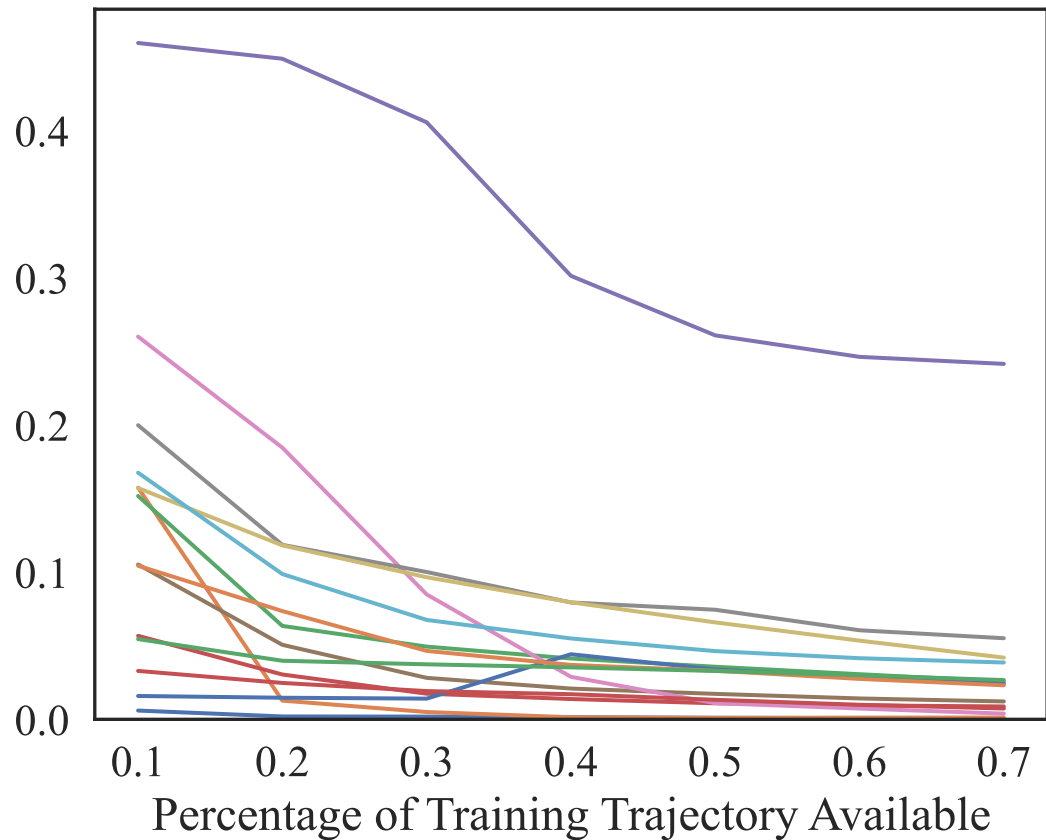


Mean Normalized Distance



original\_paper

- Revisiting Neural Scaling Laws in Language and Vision
- DataMixing arxiv.org/abs/2405.14908
- datablations
- chinchilla
- llm360-private
- LM360:arxiv-2312.06550
- training\_trajectories
- pythia
- gpt3 - arxiv.org-abs-2005.14165
- MAD arxiv.org/2403.17844
- olmo
- overtrain
- blog-redpajama-7b
- t5-pile