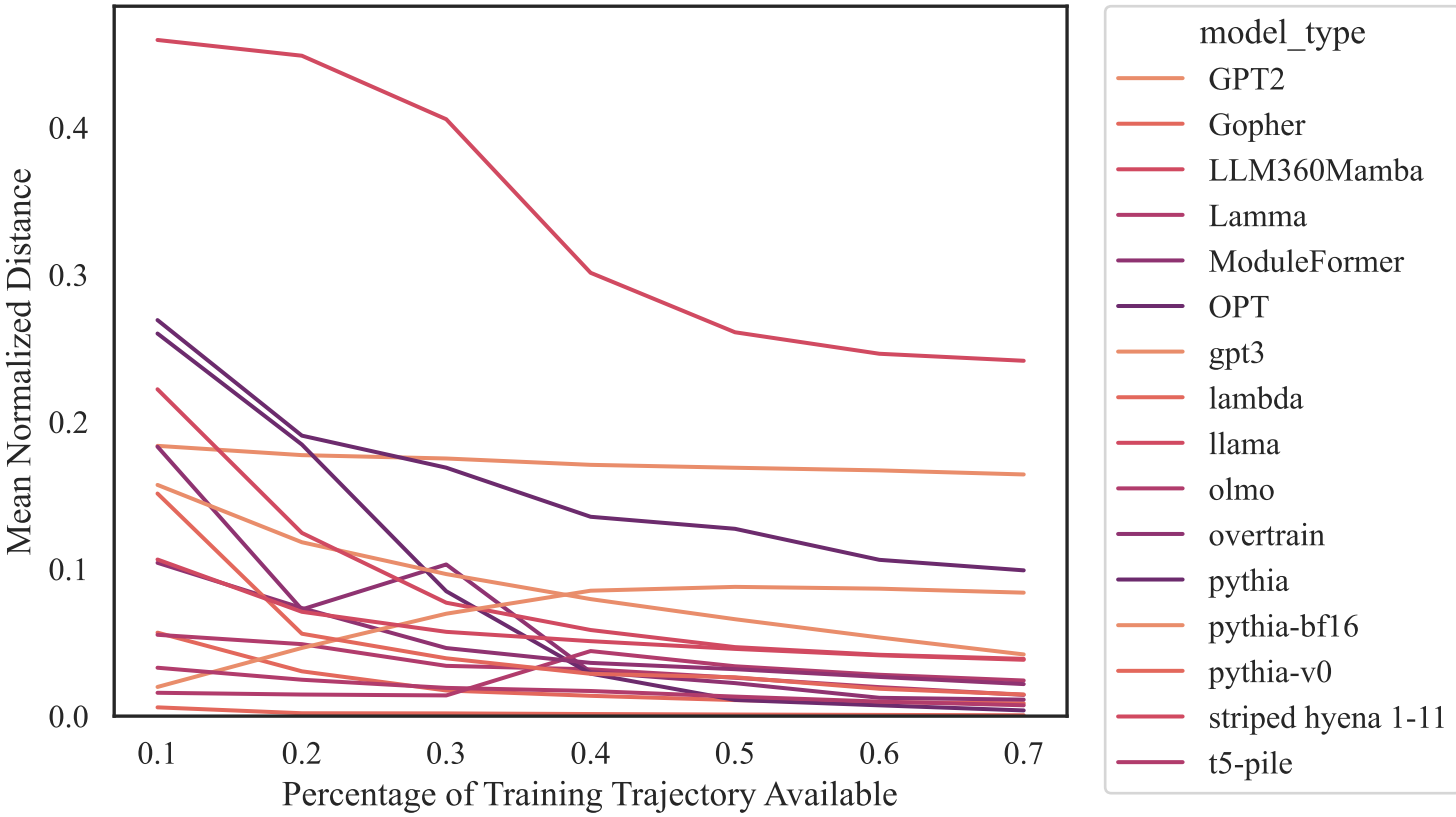
Mean Normalized Distance
0.0
0.1
0.2
0.3
0.4
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Percentage of Training Trajectory Available
model_type
GPT2
Gopher
LLM360Mamba
Lamma
ModuleFormer
OPT
gpt3
lambda
llama
olmo
overtrain
pythia
pythia-bf16
pythia-v0
striped hyena 1-11
t5-pile