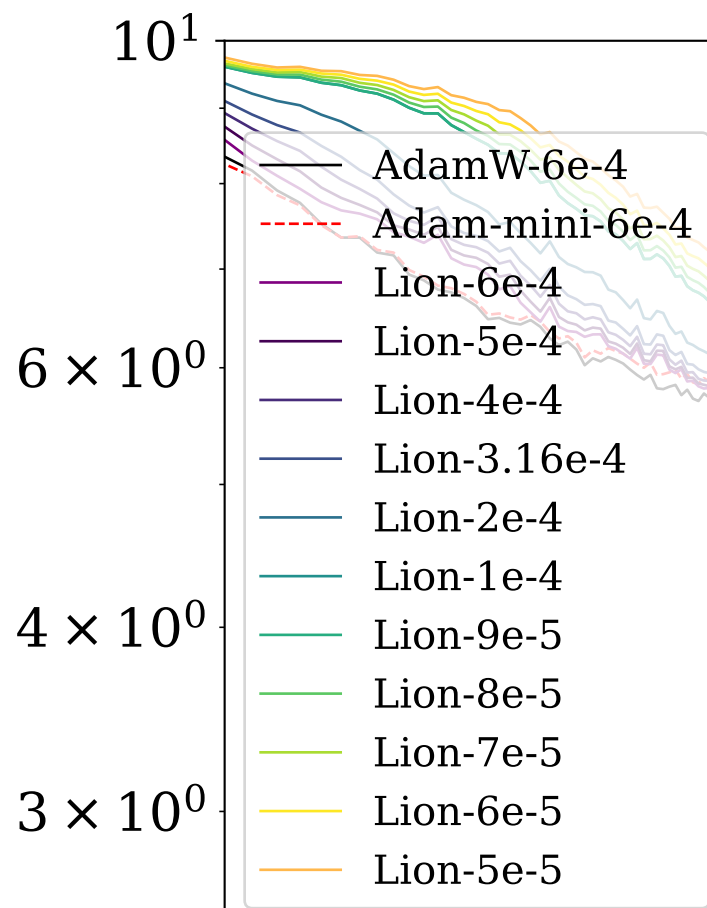


# GPT-2-125M Pre-train

Train Loss



Iteration