

CE Test Loss vs. Number of Training Tokens

