

Loss vs Time for gpt2

