

GPT-2 small, Training from scratch

