

Distribution of "pre\_attention\_norm" Weights in gpt2

