

Distribution of "post_attention_norm" Weights in gpt2

