

Distribution of "attention_output" Weights in gpt2

