Distributions for "attention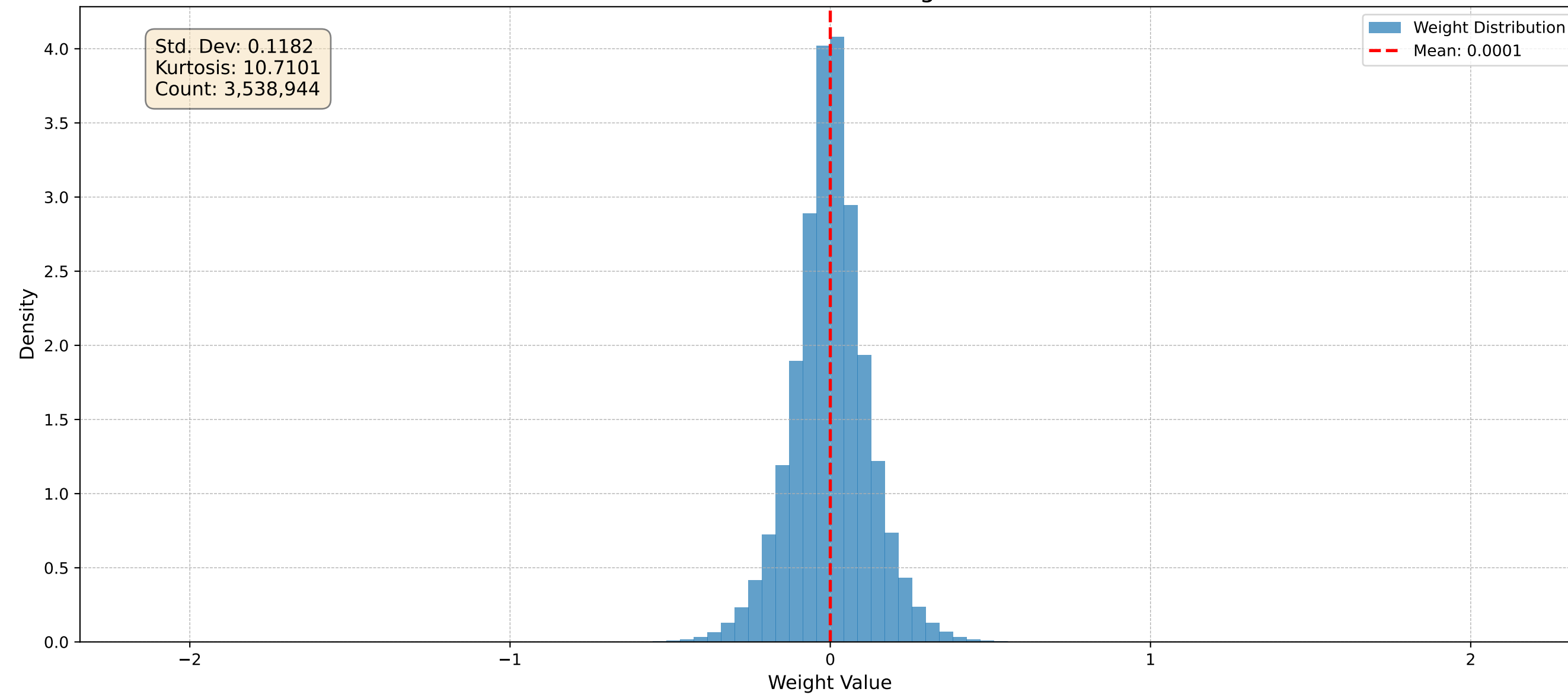_value" in distilbert/distilgpt2