gradients/bert/encoder/layer_1/attention

output
28.75%

self
71.25%