gradients/bert/encoder/layer_2

output
18.05%

attention
62.99%

intermediate
18.96%