

gradients/bert/encoder/layer_5

output
18.06%

attention
62.40%



intermediate
19.55%