

gradients/bert/encoder/layer_1

output
18.03%

attention
62.44%

intermediate
19.53%

