

gradients/bert/encoder/layer_2

output
18.04%

attention
62.73%

intermediate
19.23%

