

gradients/bert/encoder/layer_6

output
18.11%

attention
62.32%

intermediate
19.56%

