

gradients/bert/encoder/layer_5

output
18.04%

attention
62.11%

intermediate
19.85%

