

gradients/bert/encoder/layer_1

output
18.04%

attention
62.85%

intermediate
19.11%

