

gradients/bert/encoder/layer_3

output
17.87%

attention
62.82%

intermediate
19.31%

