

gradients/bert/encoder/layer_3

output
17.77%

attention
62.75%

intermediate
19.48%

