

gradients/bert/encoder/layer_3

output
17.99%

attention
62.80%

intermediate
19.21%

