

gradients/bert/encoder/layer_3

output
17.93%

attention
62.83%

intermediate
19.24%

