

gradients/bert/encoder/layer_2

output
18.26%

attention
62.20%

intermediate
19.54%

