

gradients/bert/encoder/layer_1

output
18.06%

attention
62.40%

intermediate
19.54%

