

gradients/bert/encoder/layer_2

output
22.03%

attention
42.39%

intermediate
35.57%

