

gradients/bert/encoder/layer_2

output
22.84%

attention
50.43%

intermediate
26.73%

