

gradients/bert/encoder/layer_2

output
24.88%

attention
49.72%

intermediate
25.40%

