

gradients/bert/encoder/layer_0

output
22.28%

attention
43.29%

intermediate
34.43%

