

gradients/bert/encoder/layer_4

output
20.92%

attention
45.14%

intermediate
33.95%

