

gradients/bert/encoder/layer_4

output
17.93%

attention
62.68%

intermediate
19.39%

