

gradients/bert/encoder/layer_5

output
17.79%

attention
62.68%

intermediate
19.53%

