

gradients/bert/encoder/layer_6

output
17.87%

attention
62.68%

intermediate
19.44%

