

gradients/bert/encoder/layer_4

output
17.71%

attention
63.09%

intermediate
19.21%

