

gradients/bert/encoder/layer_4

output
17.77%

attention
62.91%



intermediate
19.32%