

gradients/bert/encoder/layer_7

output
17.86%

attention
62.36%



intermediate
19.77%