

gradients/bert/encoder/layer_10

output
17.77%

attention
62.53%



intermediate
19.70%