

gradients/bert/encoder/layer_10

output
17.88%

attention
62.46%



intermediate
19.66%