

gradients/bert/encoder/layer_3

output
19.19%

attention
48.97%



intermediate
31.84%