

gradients/bert/encoder/layer_1

output
22.12%

attention
43.17%



intermediate
34.70%