

gradients/bert/encoder/layer_5

output
22.34%

attention
42.17%



intermediate
35.48%