

gradients/bert/encoder/layer_2

output
21.91%

attention
43.20%



intermediate
34.89%