

gradients/bert/encoder/layer_3

output
17.99%

attention
51.32%



intermediate
30.69%