

gradients/bert/encoder/layer_3

output
25.08%

attention
39.31%



intermediate
35.61%