

gradients/bert/encoder/layer_1

output
25.22%

attention
39.36%



intermediate
35.42%