

gradients/bert/encoder/layer_11

output
18.20%

attention
61.89%



intermediate
19.90%