

gradients/bert/encoder/layer_4

output
21.94%

attention
42.31%



intermediate
35.75%