

gradients/bert/encoder/layer_5

output
21.95%

attention
42.40%



intermediate
35.65%