

gradients/bert/encoder/layer_5

output
24.97%

attention
39.57%



intermediate
35.46%