

gradients/bert/encoder/layer_8

output
25.45%

attention
38.81%



intermediate
35.73%