

gradients/bert/encoder/layer_4

output
25.26%

attention
38.91%



intermediate
35.83%