

gradients/bert/encoder/layer_7

output
21.88%

attention
42.50%



intermediate
35.62%