

gradients/bert/encoder/layer_4

output
21.88%

attention
42.53%



intermediate
35.59%