

gradients/bert/encoder/layer_6

output
21.88%

attention
42.48%



intermediate
35.64%