

gradients/bert/encoder/layer_11

output
17.94%

attention
62.30%



intermediate
19.76%