

gradients/bert/encoder/layer_8

output
22.11%

attention
42.15%



intermediate
35.74%