

gradients/bert/encoder/layer_8

output
25.23%

attention
38.98%



intermediate
35.79%