

gradients/bert/encoder/layer_6

output
25.23%

attention
39.01%



intermediate
35.76%