

gradients/bert/encoder/layer\_3

output  
24.19%

attention  
46.50%



intermediate  
29.31%