

gradients/bert/encoder/layer\_7

output  
17.87%

attention  
62.63%



intermediate  
19.50%