

gradients/bert/encoder/layer\_9

output  
17.80%

attention  
62.63%



intermediate  
19.56%