

gradients/bert/encoder/layer\_10

output  
17.81%

attention  
62.53%



intermediate  
19.66%