

gradients/bert/encoder/layer\_2

output  
25.27%

attention  
39.05%



intermediate  
35.68%