

gradients/bert/encoder/layer\_3

output  
21.97%

attention  
42.40%



intermediate  
35.62%