

gradients/bert/encoder/layer\_1

output  
21.65%

attention  
43.25%



intermediate  
35.10%