

gradients/bert/encoder/layer\_0

output  
20.65%

attention  
49.84%



intermediate  
29.51%