

gradients/bert/encoder/layer\_0

output  
22.39%

attention  
42.52%



intermediate  
35.08%