

gradients/bert/encoder/layer\_0

output  
25.54%

attention  
39.10%



intermediate  
35.36%