

gradients/bert/encoder/layer\_8

output  
17.87%

attention  
62.56%

intermediate  
19.56%

