

gradients/bert/encoder/layer\_9

output  
17.87%

attention  
62.38%

intermediate  
19.75%

