

gradients/bert/encoder/layer\_4

output  
17.95%

attention  
62.78%

intermediate  
19.28%

