

gradients/bert/encoder/layer\_7

output  
17.82%

attention  
62.58%

intermediate  
19.60%

