

gradients/bert/encoder/layer\_0

output  
17.87%

attention  
63.31%

intermediate  
18.83%

