

gradients/bert/encoder/layer\_0

output  
17.95%

attention  
63.28%

intermediate  
18.78%

