

gradients/bert/encoder/layer\_2

output  
21.02%

attention  
52.32%

intermediate  
26.66%

