

gradients/bert/encoder/layer\_5

output  
17.97%

attention  
62.71%

intermediate  
19.32%

