

gradients/bert/encoder/layer\_5

output  
17.80%

attention  
62.46%

intermediate  
19.74%

