

gradients/bert/encoder/layer\_4

output  
22.02%

attention  
42.61%



intermediate  
35.37%