

gradients/bert/encoder/layer\_6

output  
21.80%

attention  
42.71%



intermediate  
35.49%