

gradients/bert/encoder/layer\_5

output  
21.87%

attention  
42.59%



intermediate  
35.54%