

gradients/bert/encoder/layer\_3

output  
19.30%

attention  
47.26%

intermediate  
33.45%

