

gradients/bert/encoder/layer\_4

output  
24.85%

attention  
47.40%

intermediate  
27.76%

