

gradients/bert/encoder/layer\_0

output  
25.47%

attention  
39.46%

intermediate  
35.07%

