

gradients/bert/encoder/layer_2

output
18.08%

attention
62.79%

intermediate
19.13%

