

gradients/bert/encoder/layer_4

output
17.83%

attention
62.72%

intermediate
19.45%

