

gradients/bert/encoder/layer_2

output
17.87%

attention
62.78%

intermediate
19.35%

