

gradients/bert/encoder/layer_6

output
17.73%

attention
62.60%

intermediate
19.67%

