

gradients/bert/encoder/layer_0

output
17.83%

attention
63.14%

intermediate
19.03%

