

gradients/bert/encoder/layer_8

output
17.85%

attention
62.58%

intermediate
19.57%

