

gradients/bert/encoder/layer_8

output
17.85%

attention
62.65%

intermediate
19.50%

