

gradients/bert/encoder/layer_8

output
17.76%

attention
62.53%

intermediate
19.70%

