

gradients/bert/encoder/layer_7

output
17.93%

attention
62.63%

intermediate
19.44%

