

gradients/bert/encoder/layer_6

output
17.85%

attention
62.52%

intermediate
19.63%

