

gradients/bert/encoder/layer_8

output
17.90%

attention
62.62%

intermediate
19.47%

