

gradients/bert/encoder/layer_0

output
17.90%

attention
63.14%

intermediate
18.96%

