

gradients/bert/encoder/layer_3

output
25.26%

attention
38.88%

intermediate
35.86%

