

gradients/bert/encoder/layer_1

output
24.82%

attention
42.69%



intermediate
32.49%