

gradients/bert/encoder/layer_0

output
23.98%

attention
48.19%



intermediate
27.84%