

gradients/bert/encoder/layer_2

output
25.10%

attention
39.47%



intermediate
35.43%