

gradients/bert/encoder/layer_5

output
21.86%

attention
42.41%



intermediate
35.73%