

gradients/bert/encoder/layer_8

output
25.25%

attention
39.01%



intermediate
35.74%