gradients/bert/encoder/layer_10

output
21.97%

attention
42.05%

intermediate
35.98%