gradients/bert/encoder/layer_5

output
17.89%

attention
62.78%

intermediate
19.33%