

gradients/bert/encoder/layer_8

output
25.21%

attention
39.12%

intermediate
35.68%

