

gradients/bert/encoder/layer_0

output
18.71%

attention
51.83%

intermediate
29.47%

