

gradients/bert/encoder/layer_0

output
17.89%

attention
63.10%

intermediate
19.02%

