

gradients/bert/encoder/layer_1

output
25.33%

attention
39.01%

intermediate
35.66%

