gradients/bert/encoder/layer_3
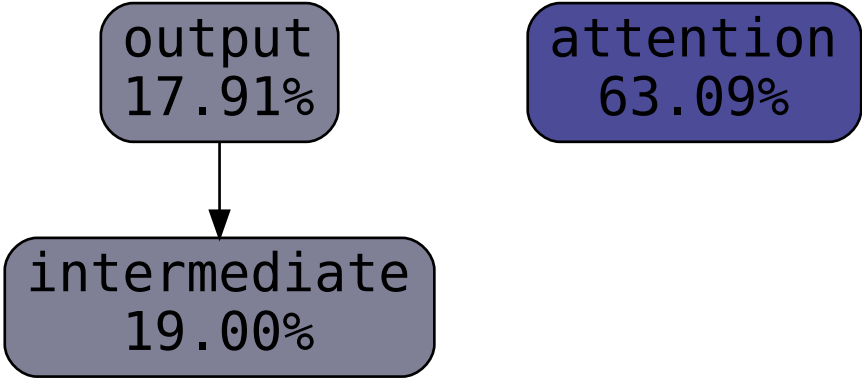
output
17.91%

attention
63.09%

intermediate
19.00%