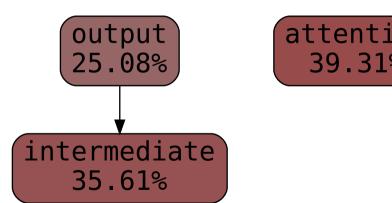
gradients/bert/encoder/layer 3



attention 39.31%