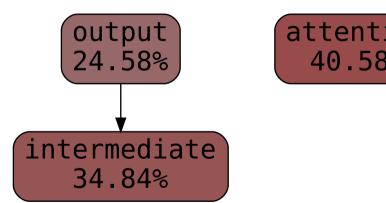
gradients/bert/encoder/layer 2



attention 40.58%