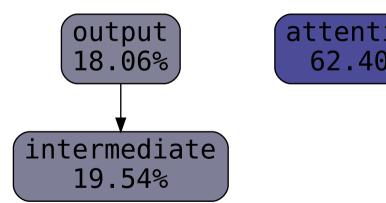
gradients/bert/encoder/layer 1



attention 62.40%