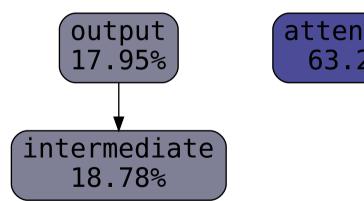
gradients/bert/encoder/layer 0



attention 63.28%