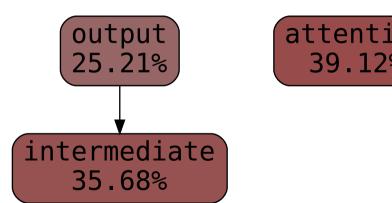
gradients/bert/encoder/layer 8



attention 39.12%