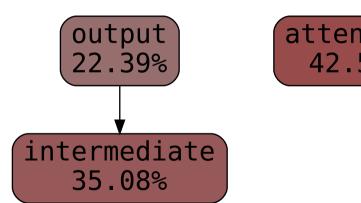
gradients/bert/encoder/layer 0



attention 42.52%