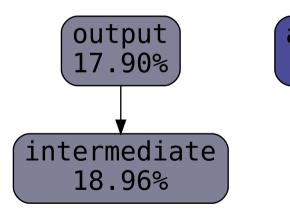
gradients/bert/encoder/layer_0



attention 63.14%