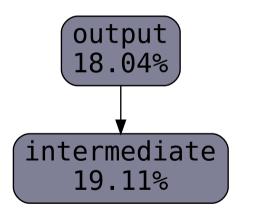
gradients/bert/encoder/layer_1



attention 62.85%