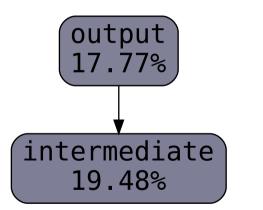
gradients/bert/encoder/layer_3



attention 62.75%