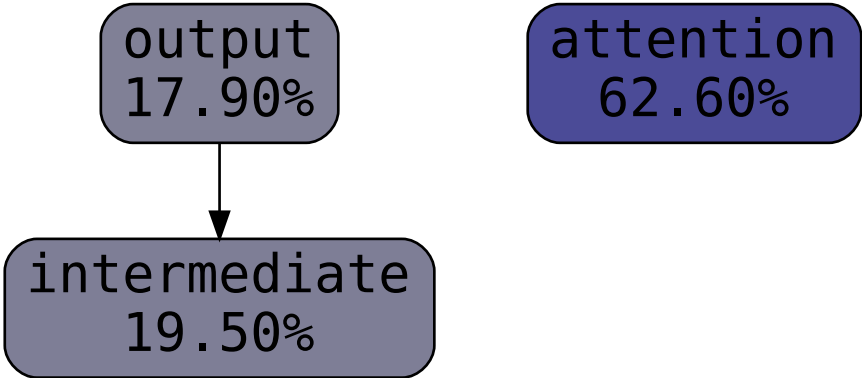gradients/bert/encoder/layer_7

output 17.90%

attention 62.60%

intermediate 19.50%