

Layer:model.layers.0.attention.wv.weight alpha=4.78, ks\_distance=0.09

