

Layer:model.layers.0.attention.wk.weight alpha=2.89, ks\_distance=0.07

