

Layer:model.layers.0.attention.wk.weight alpha=3.91, ks\_distance=0.05

