

Layer:model.layers.0.attention.wk.weight alpha=3.88, ks_distance=0.06

