

Layer:model.layers.0.attention.wk.weight alpha=3.60, ks_distance=0.07

