

Layer:model.layers.0.attention.wk.weight alpha=3.95, ks_distance=0.08

