

Layer:model.layers.0.attention.wk.weight alpha=8.20, ks_distance=0.15

