

Layer:model.layers.1.attention.wk.weight alpha=3.77, ks_distance=0.06

