

Layer:model.layers.1.attention.wk.weight alpha=3.80, ks_distance=0.05

