

Layer:model.layers.1.attention.wk.weight alpha=2.85, ks_distance=0.05

