

Layer:model.layers.0.attention.wk.weight alpha=2.81, ks_distance=0.05

